



**An anthology
of articles on
disk recording
from the pages of the
Journal of the
Audio Engineering
Society
Vol. 1-Vol. 28
(1953-1980)**

*Copyright © 1980 Audio
Engineering Society, Inc.
Library of Congress Catalog
Card No. 80-53466. Selections
from Volumes 1 through 28,
1953-1980. First printing
October 1980 (all rights
reserved). Except for brief
passages to be used in review
or as citation of authority,
no part of this book may be
reproduced without prior
permission from:
Publications Office, Audio
Engineering Society, Inc.,
60 East 42nd Street, New York,
New York 10165.*

preface

In 1893, when Emile Berliner first flattened the recording cylinder into the now familiar disk, Thomas Edison gave it little chance of succeeding. It was not until 1928 that Edison finally conceded, insuring that the cylinder would take its place among the museum oddities of the past. In a speech given at an Audio Engineering Society convention in 1960, held at the Alexandria Hotel in Los Angeles, George Brown, head of the then-existing Ampex United Stereo Tape operation, gave the phonograph record "... five years to get off the market." As a result, many manufacturers of disk-cutting equipment gave up their traditional crafts and embraced that modern miracle, tape, as "the great stringy hope." IBM changed its dictating machines to magnetic belts and the computer industry raced headlong into the consumption of reels and reels of digital tape, while consumers started a love affair with reel-to-reel, eight-track cartridges, and compact cassettes.

Those consumers of music whose habit it was to put a tall stack of records on the changer and to turn them over when all had been played, rightfully found the endless eight-track cartridge much handier. And those who enjoyed music in their automobiles, who had gone through several disappointments at the hands of 16½-r/min car record players, naturally fled to the magnetic medium.

Then came video and, again, the word was tape: easy to edit, easy to erase and reuse, easy to handle and non-wearing; but unfortunately, available in no fewer than six incompatible standards.

But the disk record would not die! Its well-engineered replication process; its large 12-inch by 12-inch point of purchase advertising area; its convenient storage; its high-storage density; the rapid access it affords without long and tedious winding; its durability and its ever-improving quality all testify to the rightness of the disk medium. Where one deals with a permanently prerecorded medium, it would seem economically odd to use magnetic tape, since one of its biggest advantages over the disk is its erasability, reusability, and almost unlimited duration of play, even, at times, at the expense of quality.

Even IBM went back to disk for its latest dictating system, while it is apparent that disk storage has revolutionized computer-memory density and access speed. The circle was further closed when special phonograph records were made by direct-to-disk techniques with telling improvement in quality.

Of course, a great deal of progress has been made since

Berliner's first disk record: progress in material selection, in plating and stamping, in turntable design, in pickup and tone-arm construction, in groove geometry, high-density cutting, and last, but not least, in the cutting lathe. The pages in this two-volume anthology tell of this progress and also of the research and development which made such progress possible. The readers will learn from these volumes how audio engineering drew on the mechanical, electrical, and material engineering sciences to achieve an outstanding system for speech and music reproduction. It should be stressed that such progress could be obtained only by international cooperation on both the engineering and marketing level.

Now the disk faces a new challenge as a storage medium for video information. The recording of video has been a spooled-ribbon medium for about as long as phonograph recording has existed. Optical sprocketed film, still the king of the motion-picture industry, soon found its match in magnetic recording, replacing one linear-motion medium with another.

But in 1976, 58 years after the first such experiments, the world's first video disk was marketed by Telefunken-Decca. Technology is now available to produce laser-recorded and -played disks, and traditionally cut and replicated disks which are playable equally on laser, electrostatic, and pressure-pickup players.

As a fall-out of this high-density video-disk technology, there are now proposals for small, very-long-playing, multi-channel audio disks, one even encased in a cassette! It is hoped that this two-volume Anthology will encourage future study and research. Volume 1 concentrates on the recording process, while Volume 2, soon to be published, deals with the playback of records. The Appendix of Volume 2 will feature a thorough patent review going back more than 100 years.

One thing is certain: the mechanical disk has developed during a 100-year period when standards could grow in an orderly fashion. Today, the long-playing record is one of the world's best standardized storage media. Never again, I feel, will any system, regardless of how simple, capture the world with a single standard equally respected in Beijing, Berlin and Baltimore.

Stephen F. Temmer

October 1980

disk recording

VOLUME 1

GROOVE GEOMETRY AND THE RECORDING PROCESS

- A Groove Geometry**
- B Disk Recording Systems**
- C Cutterheads and Lathes**
- D Styli and Lacquer Blanks**
- E Record Pressing**
- F High-Density Disk Technology**
- G Standards and Invention**
- H Related Reading**

Volume 2 DISK PLAYBACK, TESTING, AND PATENT REVIEW

to be published subsequently will contain:

**Disk Playback; Disk Pickups; Tone Arms and Turntables;
Testing and Quality Control (Records and Phonograph
Cartridges); and Patent Review.**

VOLUME 1**contents****A. GROOVE GEOMETRY**

Latest Advances in Extra Fine Groove Recording. Peter C. Goldmark (1958 July)	3
Further Thoughts on Geometric Conditions in the Cutting and Playing of Stereo Disks. C.R. Bastiaans (1963 January)	5
On "Geometric Conditions in the Cutting and Playing of Stereo Disks." Duane H. Cooper (1963 April)	15
On "Geometric Conditions in the Cutting and Playing of Stereo Disks." J. v. d. Steen (1963 July)	15
Further Comments on "Geometric Conditions in the Cutting of Stereo Records." Duane H. Cooper (1963 July)	15
A Study of Program-Level Overloading in Phonograph Recording. J.G. Woodward and E.C. Fox (1963 January)	17
Program-Level Overloading and Equalization in Phonograph Recording. E.C. Fox and J.G. Woodward (1963 January)	25
Tracing Distortion—Its Cause and Correction in Stereodisk Recording Systems. E.C. Fox and J.G. Woodward (1963 October)	31
Compensation for Tracing and Tracking Error. Duane H. Cooper (1963 October)	39
Integrated Treatment of Tracing and Tracking Error. Duane H. Cooper (1964 January)	44
Measurement of Distortions Due to Vertical Tracking Angle Errors in Stereodisk Systems. J.B. Halter and J.G. Woodward (1964 January)	50
On "Cutting-Angle Compensation." Duane H. Cooper (1964 January)	57
Further Comments on the "Measurement of Distortions in Stereodisk Systems." J.B. Halter (1964 April)	58
On Tracking and Tracing Error Measurements. Duane H. Cooper (1964 October)	61
Epilogue on Measurements. Duane H. Cooper (1964 October)	76
A New Method of Disc Recording for Reproduction with Reduced Distortion: the Tracing Simulator. Horst Redlich and Hans-Joachim Klemp (1965 April)	78
Interaction of Tracing and Tracking Error. Duane H. Cooper (1965 April)	86
Construction of Tracing Correlator Waveforms. Duane H. Cooper (1965 July)	90
Misinterpretation of Vertical Tracing Error. Duane H. Cooper (1965 July)	95
Continuous Delay Regulator for Controlling Recording Errors. Duane H. Cooper (1966 January)	97
Factors Affecting the Stylus/Groove Relationship in Phonograph Playback Systems. C.R. Bastiaans (1967 October)	107

Comments on "The Stylus/Groove Relationship in Phonograph Playback Systems." D.A. Barlow (1968 July)	118
On the Interaction between Tracing Correction and a Bandwidth Limitation. Duane H. Cooper (1969 January).....	119
Development and Application of a New Tracing Simulator. Dieter Braschoss (1971 February)	131
Groove Echo in Lacquer Masters. Daniel W. Gravereaux and Benjamin B. Bauer (1971 November)	138
Comments on "Groove Echo in Lacquer Masters." D.A. Barlow (1972 April)	142
Reply to "Comments on 'Groove Echo in Lacquer Masters.'" Daniel W. Gravereaux and Benjamin B. Bauer (1972 April)	143
Development of Skew-Sampling Compensator for Tracing Error. Shigetaka Washizawa, Tomofumi Nakatani, and Takeo Shiga (1973 October)	145
A Theory of Scanning Loss in Phonographs. James V. White (1973 March).....	150
Tracing Distortion Correction. E.G. Trendell (1977 May)	159
Groove Deformation and Distortion in Records. D.A. Barlow and G.R. Garside (1978 July/August).....	164
More About "Groove Deformation and Distortion in Records." D.A. Barlow (1979 March)	176

B. DISK RECORDING SYSTEMS

Binaural Disc Recording. Emory Cook (1953 January)	179
Monogroove Stereophonic Disk Recording. John T. Mullin (1954 October)	182
Recent Developments in Stereo Disc Recording. John G. Frayne and R.R. Davis (1959 October).....	185
Practical Aspects of High-Fidelity Disk Recording. Part 1. Carlos E.R.A. Moura (1960 October).....	192
The RCA Victor Dynagroove System. Harry F. Olson (1964 April)	203
A Brief History of the Recording Industry in the Soviet Union. A.I. Archinov (1970 February)	220
Methods of High-Quality Recording and Reproducing of Music and Speech Based on Telephone Research. J.P. Maxfield and H.C. Harrison (1978 May)	223

C. CUTTERHEADS AND LATHES

A Moving-Coil Feedback Disk Recorder. C.C. Davis (1954 October)	241
Recent Developments in Precision Master Recording Lathes. Jerry B. Minter (1956 April)	247
A New Stereo Feedback Cutterhead System. Rein Narma and Norman J. Anderson (1959 October)	253
The Westrex 3D StereoDisk System. C.S. Nelson and J.W. Stafford (1964 July)	260
A Method for Raising the Load Capability of Stereo Cutters. Horst Redlich and Hans-Joachim Klemp (1964 July).....	268

Phase-Shift Characteristics of Record Cutters and Pickups. Daniel W. Gravereaux, Arthur J. Gust, and Benjamin B. Bauer (1972 January/February)	273
--	-----

Essential Equipment for the Transmission of High Peak Levels in the Disk-Cutting System SAL 74/SX 74. O. Kern and E. Weiss (1975 November)	277
--	-----

A Real-Time Digital Processor for Disk Mastering Lathe Control. Frank Hirsch and Stephen F. Temmer (1979 January/February)	282
--	-----

D. STYLI AND LACQUER BLANKS

Practical Aspects of Hot Stylus. Carlos E.R. de A. Moura (1957 April)	289
---	-----

On the Damping of Phonograph Styli. B.B. Bauer (1964 July)	293
--	-----

Design and Use of Recording Styli. Richard Marcucci (1965 April)	297
--	-----

An Improved Disc for Master Recording. John E. Jackson (1965 April)	301
---	-----

Stylus Mass and Elliptical Points. John Walton (1966 April)	305
---	-----

Further Remarks on "Stylus Mass and Elliptical Points." John Walton (1966 April)	311
--	-----

The Diamond as an Industrial Material, With Special Reference to Phono Styli. E.J. and M.V. Marcus (1970 February)	312
--	-----

On "Diamond Styli." Paul W. Klipsch (1971 October)	316
--	-----

Biradial and Spherical Stylus Performance in a Broadcast Disc Reproducer. J.R. Sank (1970 August)	317
---	-----

Comments on "Biradial and Spherical Stylus Performance in a Broadcast Disc Reproducer." D.A. Barlow (1971 February)	322
---	-----

Reply to "Comments on 'Biradial and Spherical Stylus Performance in a Broadcast Disc Reproducer.'" J.R. Sank (1971 February)	323
--	-----

An Experimental Study of Groove Deformation in Phonograph Records. James V. White (1970 October)	324
--	-----

Lacquer Warp, Advance Ball, and Disc Cutter Dynamics. Daniel W. Gravereaux and James V. White (1976 April)	334
--	-----

Dynamic Modeling and Analysis of a Phonograph Stylus. Lawrence R. Happ (1979 January/February)	338
--	-----

E. RECORD PRESSING

Record Quality and Its Relation to Manufacturing. A.M. Max (1955 January)	351
---	-----

Control of Static Electricity on a Phonograph Record. G.P. Humfeld (1962 October)	358
---	-----

Record Contamination: Causes and Cure. Percy Wilson (1965 April)	362
--	-----

Instruments for Record Cleaning. Percy Wilson (1965 October)	370
--	-----

Deformation Distortion in Disc Records. Takeo Shiga (1966 July)	375
---	-----

Comments on "Deformation Distortion in Disc Records." Duane H. Cooper (1968 October)	385
--	-----

Performance Characteristics of the Commercial Stereo Disc. John Eargle (1969 August)	386
--	-----

A New Profile for LP Records. Warren Rex Isom (1971 July/August)	393
Long-Term Durability of Pickup Diamonds and Records. Friedrich A. Loescher and Frank H. Hirsch (1974 December)	397
Role of Polymer Science in Developing Materials for Phonograph Discs. S.K. Khanna (1976 July/August)	401
Record Warps and System Playback Performance. Larry Happ and Frank Karlov (1976 October)	407

F. HIGH-DENSITY DISK TECHNOLOGY

Phonovid—A System for Recording Television Pictures on Phonograph Records. Kenneth E. Farr (1968 April)	419
On "Recording Television Pictures on Phonograph Records." Albert Abramson (1968 October)	423
Reply to "On 'Recording Television Pictures on Phonograph Records.'" T.S. Cole (1968 October)	423
High-Density Disc Recording Systems. Leo M. Levens (1969 June)	424
A Mechanical Disc Recording and Reproducing System with High Storage Density and High Rate of Transmission. Gerhard Dickopp, Hans-Joachim Klemp, Horst Redlich, and Eduard Schueler (1970 December)	427
A Long-Play Digital Audio Disk System. Toshi Tada Doi, Takashi Itoh, and Hiroshi Ogawa (1979 December)	433
A New PCM Audio Disk Pickup Employing a Laser Diode. K. Okada, T. Kubo, W. Susaki, and T. Sato (1980 June)	440
The MD (Minidisk) System: A Contribution to the Digital Audio Disk Standard. Klaus Welland and Horst Redlich (1980 July/August)	444

G. STANDARDS AND INVENTION

AES Standard Playback Curve. (Audio Engineering Magazine, 1951 January)	451
British Patent Specification 394,325. Alan Dower Blumlein (1958 April)	453
A Report on the Proposed NAB Disc Playback Standard. John J. Bubbers (1964 January)	462
The RIAA Engineering Committee. H.E. Roys (1968 January)	466
Weighted Peak Flutter Measurement—A Summary of the New IEEE Standard. J.G. McKnight (1971 November)	469
The Benefits and Dangers of Standardization. Stephen F. Temmer (1979 May) .	472
On RIAA Equalization Networks. Stanley P. Lipshitz (1979 June)	473
Digital Audio Disk Standardization Conference. (1980 July/August)	497

H. RELATED READING

Bibliography of Disk Recording. Alfred Jorysz (1954 April)	503
Quadraphonic Disk Recording Bibliography	520

groove geometry

A



Latest Advances in Extra Fine Groove Recording*

PETER C. GOLDMARK†

CBS Laboratories, Stamford, Connecticut

Since the development of the seven-inch 16 $\frac{2}{3}$ rpm extra fine groove record, further investigations have been carried out in this field. It has been found possible to double the amount of spoken word information that could be stored on that size record.

Some of the underlying theories will be presented and demonstrations will show how near the theoretical limit current results are.

Significant portions of this work have been carried out under the auspices of the American Foundation for the Blind.

SINCE THE end of 1945 the CBS Laboratories' work in the recording field has been chiefly concentrated towards increasing the amount of information that can be recorded per unit area. With the development of the long playing record, an improvement in quality over the previous record, the now more or less defunct 78 rpm, has also been brought about. In the subsequent records, such as the 16 rpm 7-inch extra fine groove record, no improvements in quality over the LP record were attempted. A new record development reported here is directed primarily toward a field different from music, though not necessarily excluded from it.

The wavelength λ traced at any given point of the record is V/f ; where V is the linear velocity and f is the frequency of the recorded signal. For proper reproduction tracing distortion has to be a minimum; thus the effective radius of the reproducing stylus should be no bigger than the smallest radius of the recorded wave.

The effective radius is approximately $r/\sqrt{2}$, where r is the stylus tip radius, and the minimum radius of curvature of a sine wave is $\lambda^2/4\pi^2 D$, where D represents the peak-to-peak displacement of the groove.^{1,2}

Thus to minimize tracing distortion,

$$r \leq \frac{\lambda^2 \sqrt{2}}{4\pi^2 D} = \frac{V^2 \sqrt{2}}{4\pi^2 f^2 D}. \quad (1)$$

If the cutting stylus has a perfect point and perfectly

sharp edges, one can record any frequency—including video—provided the reproducing stylus tip is small enough or the recording amplitude is low enough.

In order to establish the maximum density at which information can be stored on a record, certain practical limitations have to be taken into account.

The cutting stylus will rarely come to a perfect point; thus the bottom of the resultant groove will have a finite radius. The reproducing stylus tip radius must be of finite size in order to avoid scraping the bottom of the groove, which, as pointed out above, does have a finite radius. Also, since the stylus force cannot be infinitely small, there again has to be a lower limit to the radius of the stylus tip; otherwise the record groove would be deformed if the walls could not support the stylus.

Practical records and reproducers were made with essentially the same performance characteristics as found on the inside of LP records, using reproducing needles of 0.00025-inch radius at 2 grams force with a linear velocity of 3.5 ips at the inside of the record with 550 lines per inch.

Let us now apply Eq. 1 to four different records and find the maximum value of groove modulation at which the stylus tip effective radius equals the smallest radius of the recorded wave at frequency f . Solving (1) for D ,

$$D = \frac{\sqrt{2}}{r} \left(\frac{V}{2\pi f} \right)^2 \quad (2)$$

Assuming for all four records that $f = 10,000$ cycles and V will be the linear velocity of the innermost grooves on the record, the following values for D will result (Table I).

The values for D (groove modulation) in the table are the measure of performance in the various records when using the same frequency f .

In order to maintain constant signal-to-noise ratio on all records, the deviations in modulation value of the various

* Original version delivered before the Ninth Annual Convention of the Audio Engineering Society, New York, October 8, 1957.

† President and Director of Research.

¹ P. C. Goldmark, R. Snevangers and W. S. Bachman, "The Columbia Long-Playing Microgroove Recording System," *Proceedings of the I.R.E.*, Vol. 37, August, 1949.

² P. C. Goldmark, "Highway Hi-Fi," *Audio*, Vol. 39, December, 1955.

TABLE I. Maximum values of groove modulation for various types of records. $f = 10,000$ cps.

Record	r (mils)	V (ips)	Playing time/side	Groove/inch	D (inches)
78 rpm	3	16.3	4 min	80	3.1×10^{-5}
LP (33.3 rpm)	1	9.6	25 min average	230	3.2×10^{-5}
XLP (16.6 rpm)	0.25	3.5	music: 45 min speech: 1 hr	550	1.7×10^{-5}
XXLP (8.3 rpm)	0.25	1.75	music: 1½ hr	600	0.43×10^{-5}
XXLP (8.3 rpm)*	0.25	1.75	speech: 2 hr	600	1.7×10^{-5}

* $f = 5,000$ cycles.

records have to be compensated for by decreased noise figures.³

In the case of the XLP records (for the Automobile Phonograph) this is accomplished partially by the reduced linear speed and partly by pickup design. Thus the 7-inch XLP record turning at 16.6 rpm is capable of giving a performance equal to that of the LP record when carrying twice as much playing time as the latter.

³ Tests have shown that when reducing record speed from 33 rpm to 16 rpm the average surface noise will be decreased by 4 db. Going from 16 rpm to 8 rpm, the surface noise is reduced by another 3 db. These tests were made at a constant groove diameter.

For speech, the XXLP record turning at 8.3 rpm permits the same amount of groove amplitude as the XLP record, but is limited to one half of its upper frequency. But for speech, 5000 cycles maximum frequency will give adequate reproduction.

Development of the 8 rpm fine groove record was carried out under the sponsorship of the American Foundation for the Blind, whose efforts in this field were supported by the Library of Congress. The primary purpose for the development of this new record was to investigate the possibility of providing maximum listening time per record to the blind. A single 8 rpm record of 7-inch diameter provides 4 hours of speech; the 12-inch diameter about 10 hours.

The tone arm used to reproduce these records is the same as the one used on the Automobile Phonographs. With a stylus force of 2 grams, it is able to track the extra fine grooves reliably, even in the presence of shock and vibration. In addition, it is almost impossible to scratch the records when using this type of tone arm, which is of special importance to the blind users.

During the development work just described, the painstaking and enthusiastic assistance of Thomas A. Broderick and Wilbur W. Clade of the CBS Laboratories' staff is gratefully acknowledged.

Further Thoughts on Geometric Conditions in the Cutting and Playing of Stereo Disks*

C. R. BASTIAANS

Philips Phonographic Industries, Baarn, Netherlands

The theoretical considerations presented in an earlier paper¹ are subjected to further practical analysis. The influences of vertical tracking error, in particular, are investigated (differences in slant). It is established that at high frequencies, where distortion is most annoying, the resulting *frequency-independent tracking* distortion is overpowered by *frequency-dependent tracing* distortion. The standardization of a "vertical" cutting angle, which has been suggested by others, is found to be objectionable. It is shown that slanted cutting (artificial cutting angle) limits the allowable recording level. A 10° vertical angle already limits the level by 3 db.

A new method of measurement of the separation of stereo pickups is presented. The influence of recorded crosstalk in a normal stereo test record is thus evaded.

This method also allows the determination of the orientation axes of the pickup.

The advantages of a direct 45/45° cutter compared to an indirect cut by a matrixed 0/90° cutter are shown.

I. THE INFLUENCE OF VERTICAL TRACKING ERROR

1.1 Introduction

If the angles of cutting and playback planes relative to the normal to the disk surface do not coincide, i.e., a vertical tracking error prevails, the following detrimental effects may be expected: a. *loss of separation*; b. *linear distortion*; and c. *non-linear distortion*. These effects will be investigated in the following paragraphs.

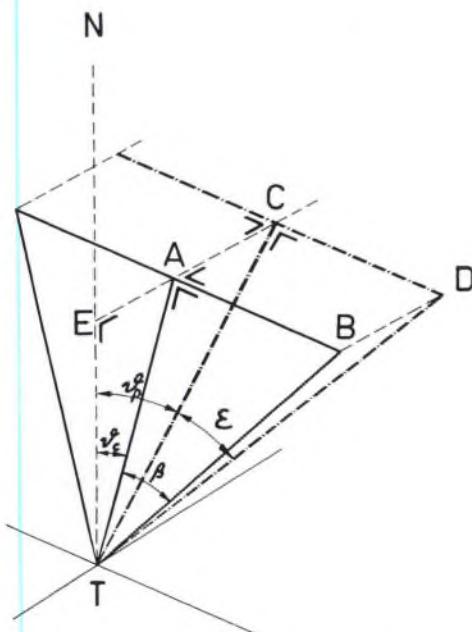


FIG. 1. Perspective view of slanted axes array (TAB for cutting plane and TCD for playback plane).

* Presented October 16, 1962 at the Fourteenth Annual Fall Convention of the Audio Engineering Society, New York.

¹ J. L. Ooms and C. R. Bastiaans, *J. Audio Eng. Soc.* 7, 115 (1959).

1.2 Theoretical considerations

1.2.1 *Loss of separation*.—Consider Fig. 1, where the cutting plane TAB and the playback plane TCD are shown slanted relative to the normal to the disk surface NT. The cutting plane is slanted at an angle θ_c , the playback plane at an angle θ_p .

Presuming the axes-orientation of the cutter-head to be correct, the angle $\beta = 45^\circ$.

We can then write (see Appendix A):

$$\cot \epsilon = \cos \theta_c / \cos \theta_p \quad (1)$$

where ϵ is the angle of axis orientation for true playback (infinite separation).

If cutting has taken place in a perfectly vertical plane, i.e., $\theta_c = 0^\circ$, this equation takes the form of

$$\tan \epsilon = \cos \theta_p. \quad (2)$$

If the axes-orientation of the pickup is chosen so as to satisfy these equations, no deterioration of the originally

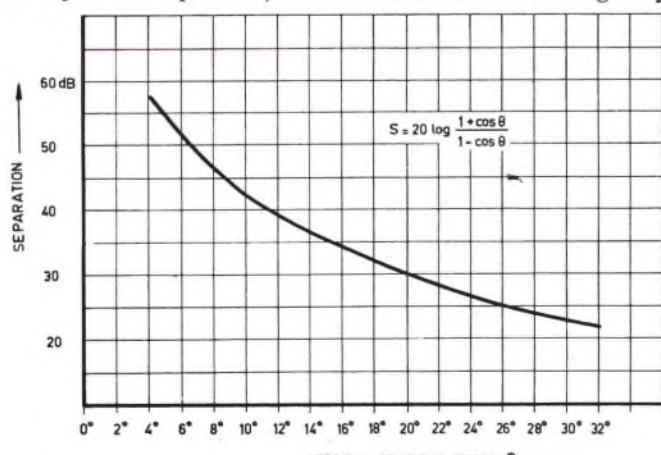


FIG. 2. Decrease of ideal separation due to vertical tracking error.

recorded crosstalk takes place. Without such a compensation additional crosstalk results; the relation between theoretical channel separation and the vertical tracking error Θ is illustrated in Fig. 2.

It can be seen that for values of $\Theta < 15^\circ$ the crosstalk due to non-coincident slant is negligible, as already established.¹

I.2.2 *Linear distortion*.—Figure 3 depicts the situation

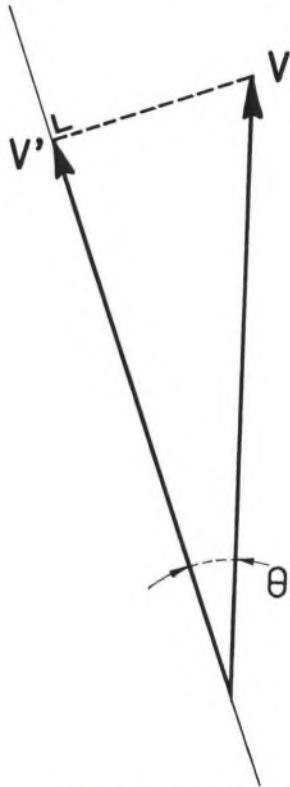


FIG. 3. Vector presentation of amplitude loss due to vertical tracking error.

when a vertical signal (vector V) is traced along an axis differing an angle Θ from the vertical. The playback level of the vertical signal is represented by the projection V' . The amplitude loss is a cosine function of the tracking error Θ , shown in Fig. 4. From the graph it may be established that the amplitude loss due to tracking error is of little importance, error angles up to 25° resulting in a signal loss of less than 1 db.

I.2.3 *Non-linear distortion*.—Appendix B gives the derivation of the following distortion formulae:

$$D_2 = (v/V) (\cos \theta_c \tan \theta_p - \sin \theta_c) \times 100\% \text{ if } \theta_p > \theta_c \quad (3)$$

$$D_2 = (v/V) (\sin \theta_c - \cos \theta_c \tan \theta_p) \times 100\% \text{ if } \theta_p < \theta_c \quad (4)$$

$$D_2 = (v/V) \tan \theta_p \times 100\% \text{ if } \theta_c = 0^\circ \quad (5)$$

$$D_2 = (v/V) \sin \theta_c \times 100\% \text{ if } \theta_p = 0^\circ. \quad (6)$$

It is seen that the *tracking* distortion is a wavelength effect and only dependent on cutting velocity v , groove ve-

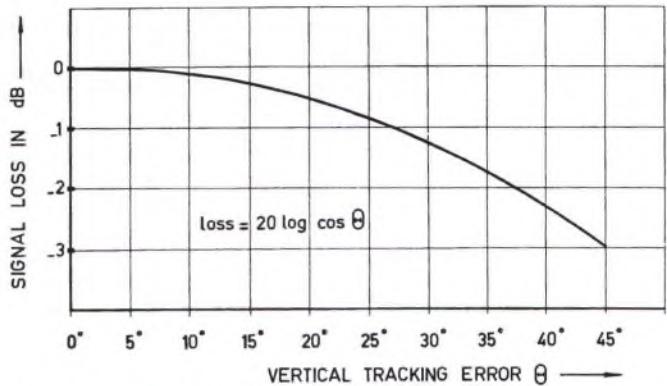


FIG. 4. Amplitude loss as a function of vertical tracking error.

locity V and tracking error ($\theta_p - \theta_c$), while it is independent of frequency. *Tracing* distortion (due to finite size of the spherical needle tip), however, is dependent on frequency as well, increasing linearly with it and also inversely with the square of the groove velocity. It may therefore be expected that for the higher frequencies (where distortion is more likely to be annoying) the tracking distortion is of minor importance as compared to the tracing distortion.

I.3 Theory versus practice

I.3.1 *Loss of separation*.—Further on in this paper, a method is presented to measure the separation of a pickup *independent of recorded crosstalk*. Using this method an infinite value of separation is possible to measure, to wit when the orientation of the axes is exactly 45/45.

With a specially designed 2-dimensional manipulator built in a tonearm, a pickup cartridge could be made to slant over a range of angles. Optical measurements of the vertical angle θ_p of the pickup used showed the value of θ_p in normal cartridge mounting to be 27° .

Variation of slant took place between the limits of -8° and $+10^\circ$, meaning an effective variation of θ_p from 19° to 37° . Since the 1000-cps test record used was made with a Westrex 3C cutter of which the vertical cutting angle θ_c is known to be 23° , a slant of -4° of this particular pickup would mean a *zero* tracking error. The measured separation results at 1000 cps are shown in Fig. 5 and compared with the theoretical predictions using Eq. (1).

Although a *quantitative* comparison between curves does not show full agreement, a *qualitative* comparison gives satisfactory results. Infinite separation occurs at a slant angle variation of about -4° , as predicted by theory. For the remainder, it may be established that separation for practical tracking errors remains adequate.

I.3.2 *Linear distortion*.—Various measurements have been done with slant angle variations up to 30° , showing level differences at 1000 cps of less than 1 db when compared with the level obtained when the pickup angle equals the cutting angle. The agreement with the theory is entirely satisfactory (refer to Fig. 4).

I.3.3 *Non-linear distortion*.—By playback of a special 1000-cps test record, cut with a known cutting angle of 23° , and varying vertical playback angles, second harmonic dis-

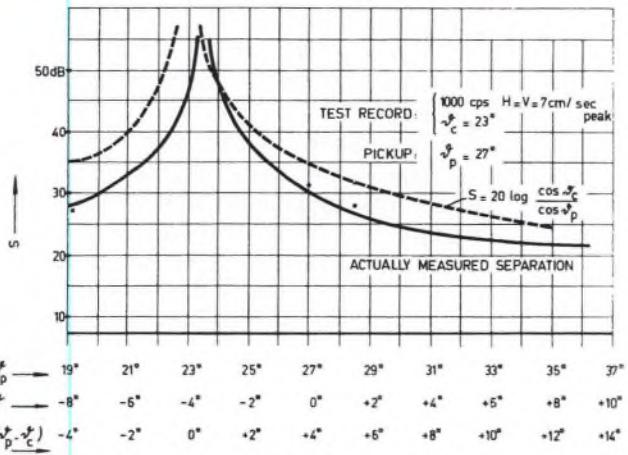


FIG. 5. Variation of separation due to vertical tracking error, compared with theoretical predictions.

tortion has been measured and plotted as a function of the slant variation. It was found that (for the type of pickup used) minimum distortion occurred with an angle variation of -5° relative to the normal mounting position of the pickup cartridge. It should therefore be expected that the tracking angle of this pickup is $23 + 5 = 28^\circ$. Optical measurement of the tracking angle with the cartridge in normal position proved to give a value of about 27° , which is a very satisfactory agreement with the theoretical value.

The measured distortion values are shown in Fig. 6 as a function of tracking error θ , where $\theta = 0^\circ$ at a slant angle variation of -5° . Formula (3) of Paragraph I.2.3 is also drawn graphically in this figure. As is to be expected, the measured second harmonic distortion is higher than the theoretical values because the measured output contains non-separable second-order products of tracing distortion as well. The slope of both curves is about identical; the measured curve slopes down to a residual value given by tracing distortion alone (calculated value of 2.5%). The agreement with the theory is quite satisfactory. In practice, however, we are faced with total harmonic distortion. A relatively large number of stereo test records of various

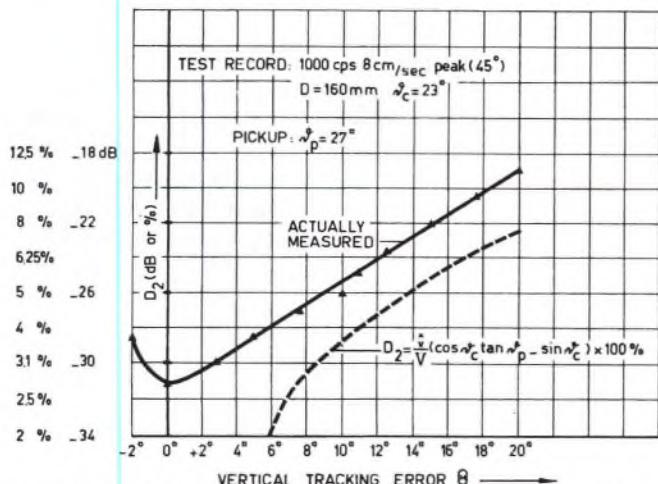


FIG. 6. Second harmonic distortion as a function of vertical tracking error, compared with theoretical curve.

makes has been traced with three types of pickups, one having a tracking angle of 37° , the second one of 34° and the third of 0° . From the stereo test records only the 1000 cps bands were selected, and then only the left-hand channel. It has been assumed that these records could be grouped in batches where $\theta_c = 23^\circ$ and $\theta_c = 0^\circ$.

The measured *total* harmonic distortion values (only the fundamental frequency removed) were plotted as a function of v/V (Figs. 7, 8 and 9) and compared with the theoretical

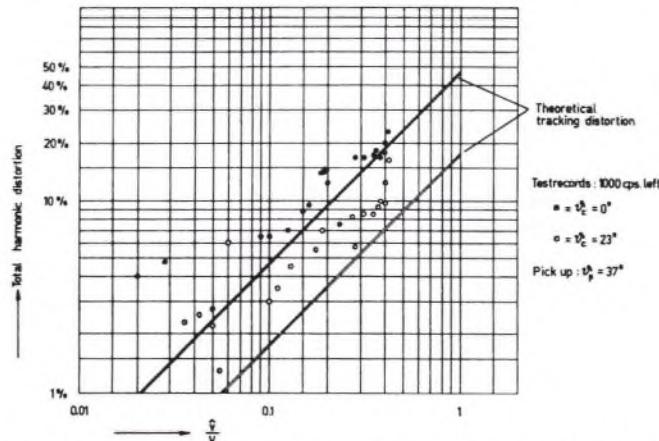


FIG. 7. Actually measured total harmonic distortion as found with various 1000-cps test records, compared with calculated second-harmonic tracking distortion; vertical tracking angle of pickup 37° .

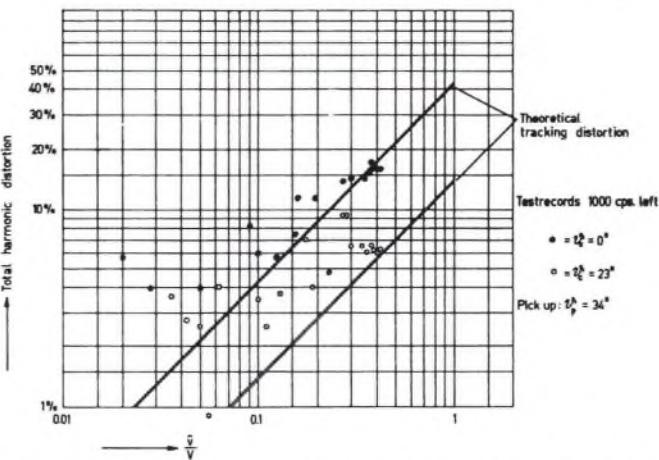


FIG. 8. Actually measured total harmonic distortion as found with various 1000-cps test records, compared with calculated second-harmonic tracking distortion; vertical tracking angle of pickup 34° .

values of *second-order tracking* distortion. It proves that the measured values, which are again the *combined* results of total tracking and tracing distortion products, are mostly higher than the theoretical values for tracking distortion alone, but the trend of the slope is comparable with the theoretical prognosis.

A further investigation was deemed necessary. By playing a 1000-cps test record in counter-clockwise rotation and tracing inner wall modulations with the right-hand channel of the pickup (tone-arm with negative off-set angle, set beyond the center spindle!), measuring results of total harmonic distortion are obtained based upon a very much

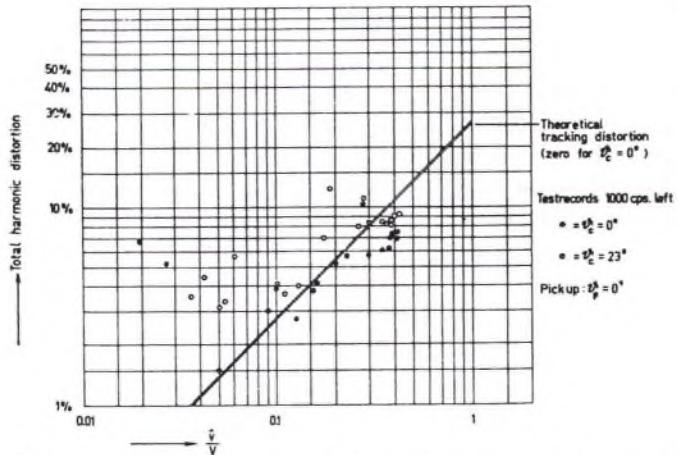


FIG. 9. Actually measured total harmonic distortion as found with various 1000-cps test records, compared with calculated second-harmonic tracking distortion; zero vertical tracking angle of pickup.

increased tracking error $\Theta = (\theta_p + \theta_c)$. These results can be compared with measurements of the same inner-wall modulation traced in the normal way, thus introducing a tracking error $\Theta = (\theta_p - \theta_c)$. In this way one may rest assured about the equality of various parameters. The only difference is the use of opposite pickup-systems, but it is felt that this does not contribute to distortion. The usability of this procedure has been proved by means of recordings with $\theta_c = 0^\circ$, from which equal distortion values are obtained with either method of playback.

Various measured results are tabulated in the following table:

Calculated D_2					
Tracking		(Tracking)	(Tracing)	Measured D_2	
θ_c	θ_p	$\Theta = 11^\circ$	5.6%	4.7%	6.5%
23°	34°	$\Theta = 57^\circ$	27.6%	4.7%	15.5%
23°	0°	$\Theta = 23^\circ$	11 %	4.7%	9.5%
-23°	0°	$\Theta = 23^\circ$	11 %	4.7%	9.5%
0°	0°	$\Theta = 0^\circ$	—	4.7%	7.5%

As can be seen, the theoretical values do not always agree with the practical results. Obviously the total harmonic distortion produced by the pickup is not necessarily an addi-

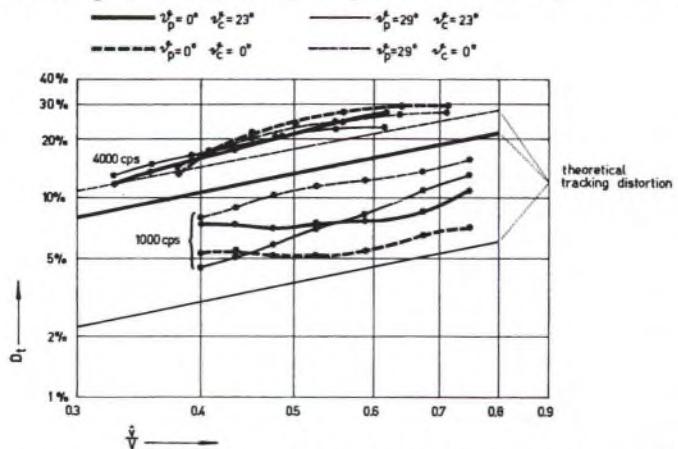


FIG. 10. Actually measured total harmonic distortion of 1000- and 4000-cps test bands compared with calculated second-harmonic tracking distortion; vertical tracking angles of pickup 0° and 29°.

tion of tracking and tracing distortion. It might even be assumed that tracking distortion may sometimes counteract tracing distortion! Now all these measurements were done with 1000-cps bands. Since distortion is generally more annoying at higher frequencies, further tests were done at 4000 cps. Special 1000- and 4000-cps test records were cut with $\theta_c = 0^\circ$ and $\theta_c = 23^\circ$ and played back with 2 types of pickups, one having a $\theta_p = 0^\circ$ and the other a $\theta_p = 29^\circ$.

The 2×4 series of measured total harmonic distortion values are shown in Fig. 10, together with theoretical distortion curves due to tracking error. It is clearly shown that at 4000 cps substantially no differences are found in measured total distortion; in other words, *tracing distortion is dominating* (calculated D_2 values amount to 11% for $v/V = 0.35$ and 31.5% for $v/V = 0.6$).

Note: When one-wall modulations are concerned, it should be remembered that all vertical angles are to be recalculated towards projections in the 45° -plane, according to the formula:

$$\sin \Theta_{45} = \sin \Theta_{90} / \sqrt{2} \quad (7)$$

where Θ_{45} is the projection of the vertical angle Θ_{90} in the 45° -plane (included angle of the groove taken to be 90° and symmetric).

1.4 Conclusion

In the light of the investigations it may be established that the effects of vertical tracking error are not as serious as is sometimes assumed.

Nevertheless, it is yet to be recommended that a *fictive* cutting angle be standardized, in order to enable the pickup manufacturers to design their products towards a compensation of one and the same "tracking error." An arbitrary value of 12° seems to be appropriate.

II. DETRIMENTAL EFFECTS OF SLANTED CUTTING

II.1 Artificial slant of cutters with zero angle

It has been suggested to standardize a certain slanted cutting angle, for instance 15° . Stereophonic cutting heads able of performing a purely vertical movement should then be tilted forward.

This now may detrimentally affect the maximum allowable cutting velocity. The back-angle of the cutting stylus is artificially increased by an amount equal to the tilting angle. This situation is depicted in Fig. 11. It is obvious that beyond a certain value of modulation depth embossing of the groove walls by the clearance faces (back planes) of the stylus takes place. Heavy flats will be caused in the otherwise smooth groovectcut, leading to very serious second-order distortion.

The maximum vertical recording velocity is given by:

$$v_{max} = V \cot(\phi + \theta_c) \quad (8)$$

where V stands for the groove velocity, ϕ for the stylus back angle and θ_c for the (artificial) slant angle of cut (see Appendix C). Normally the back angle ϕ amounts to 45° , meaning that the maximum recording velocity equals the

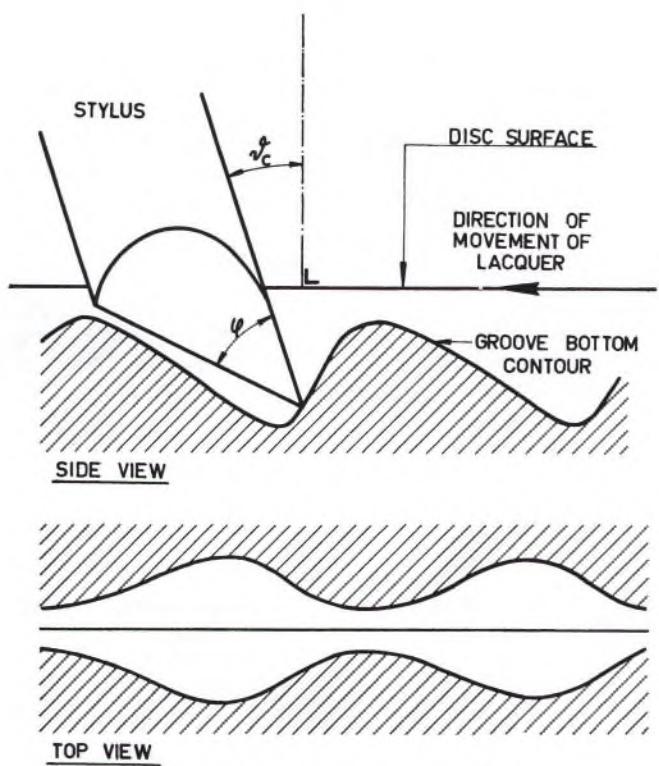


FIG. 11. Groove cutting with slanted cutting stylus.

groove velocity if $\theta_c = 0^\circ$. Any introduction of an artificial slant angle, however, decreases this maximum value. The extent to which this occurs is illustrated in Fig. 12. An artificial slant of 15° reduces the maximum modulation depth by almost 5 db! At a recording diameter of 120 mm (4 3/4 in.) such a slant limits the stylus velocity to about 8 cm/sec per channel ($33\frac{1}{3}$ rpm). Exceeding this maximum value will lead to embossing of groove walls. The groove pictures of Fig. 13 show the nature of this embossing as compared with a groove cut with the same modulation depth at a larger recording diameter. The foregoing proves that slant angle cutting with a normal type of stylus may lead to a serious limiting of levels.

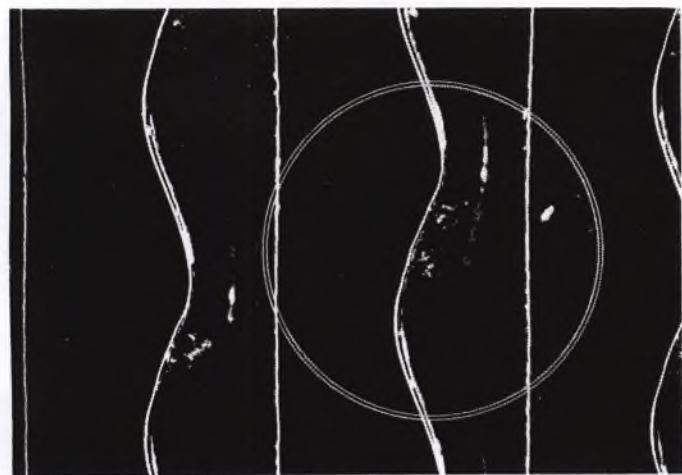


FIG. 13. Left: embossing of slanted-cut groove wall due to clearance faces of stylus touching the groove. Right: Same modulation

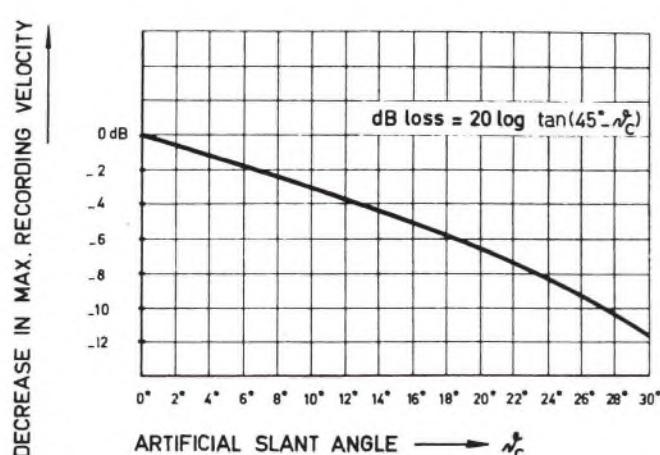


FIG. 12. Limitations set to maximum allowable recording velocity, due to slanted cutting with normal cutting styli.

Another disadvantage of slant angle recording is that the tendency of the stylus to be dragged into the lacquer requires extra measures to assure constancy in the depth of cut and avoid inadvertent variation of silent groove width.

One solution to overcome these drawbacks is to decrease the back angle of the cutting stylus. Figure 14 shows a side view of styli having different back angles accommodating slant angles up to 30° . The main disadvantage of such special styli is that the taper for compensation of slants beyond 10° is so very much increased as to cause the stylus to become fragile and easy to chip, the more so because it is "digging" into the lacquer. To overcome this disadvantage the cutting face (front plane) of the stylus could be kept at an angle of 45° with the back edge, and an angle θ_c with the stylus axis as shown in Fig. 15.

The use of such very special-purpose styli, not usable for normal monophonic recording heads, is not attractive. Moreover the proper alignment of such special styli in a cutterhead is difficult to perform.

II.2 Problems with cutters with inherent vertical "error"

Returning to the question whether a standardized cutting angle is practical, it should be borne in mind that apart from



level, slanted-cut at a higher groove speed, without embossing effect.

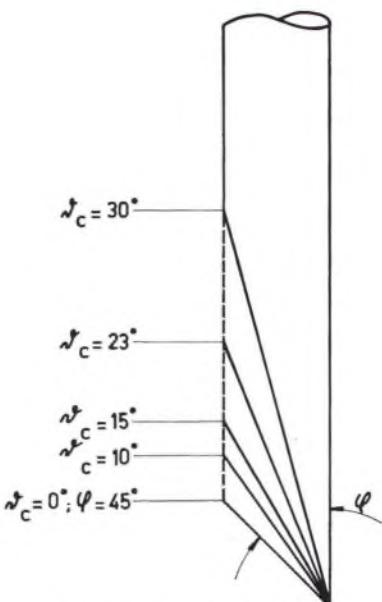


FIG. 14. Simplified side view of cutting stylus with different reductions of back angle for slanted cutting.

the above described difficulties with stereo cutter heads with pure vertical movements, cutter heads with inherent slant angles, like for instance a well-known type with $\theta_c = 23^\circ$, still offer the problem of how to realize such a standardized cutting angle. This particular cutter can not be tilted backward for more than about 4° , due to obstructions in the mechanical structure of the head, meaning that the inherent cutting angle of 23° cannot be reduced below 19° .

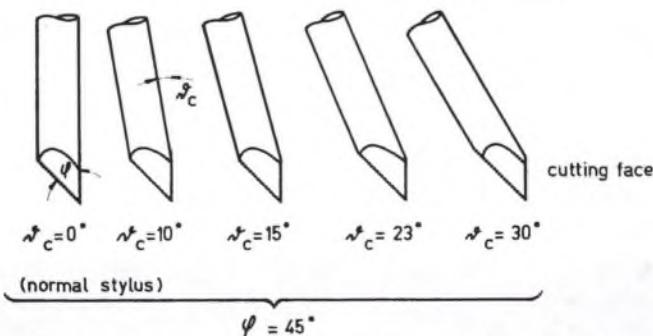


FIG. 15. Side views of special cutting styli of a form adapted to overcome disadvantages of the type shown in Fig. 14.

II.3 Conclusion

A standardized slant angle of cut with contemporary cutter heads is not practically possible to realize. In the light of the conclusions reached in Chapter I it may however be established that the phenomenon "vertical tracking error" is not as serious of character as sometimes is assumed.

III. MEASUREMENT OF SEPARATION OF STEREO PICKUPS

III.1 Shortcomings of present measuring method

Separation is normally measured by means of a test record having left and right channel modulations with mostly

unknown levels of recorded crosstalk. Even if crosstalk levels as present in the groove are known, the *phase relationship* between crosstalk and main signal remains in the dark. The crosstalk generated in the pickup system has a definite, frequency-dependent and probably amplitude-dependent phase relationship with the main signal output as well. How much this may differ between various brands and types of pickups is illustrated in Fig. 16.

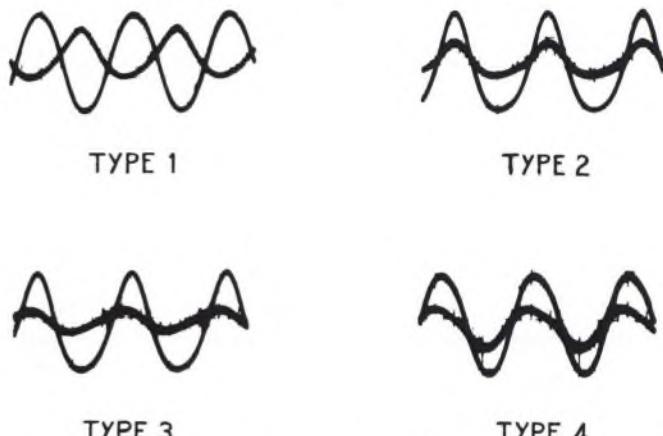


FIG. 16. Oscillograms of phase relationship between crosstalk and main signal as produced by 4 different types of pickups tracing the same test record (1000 cps). (Amplitude relation is of no significance.)

Here oscillograms of the output signals of either channel are shown, indicating the phase angle between main signal and crosstalk. Four types of pickups are used, playing the *same band* of the *same test record*. The oscillographed amplitude relation between main signal and crosstalk is of no significance. Higher harmonics are suppressed, so that the oscillograms show signals of mainly sinusoidal character. It can be seen that the phase angle varies from 145° (lagging) with pickup-type 1, about 0° with type 2 and about 35° (leading) with types 3 and 4.

From these oscillograms it may be established that the crosstalk vector as caused mechanically in the pickup system not only varies in magnitude but in direction as well. As a total result, the generated crosstalk may partly *compensate* the recorded crosstalk, and a *higher* value of separation may be measured than has ever been recorded in the test record. Quite obviously, this commonly used method of measuring the separation of a stereo pickup is far from exact. This fact has been proved in a different way. By feeding a small signal into the idle recording channel while recording a given signal in the opposite channel, it is possible to establish the phase relationship between the recorded signal and the recorded crosstalk. If the additional compensating signal results in a greater width of light pattern image reflected by the "idle" groove wall, it is obviously added to the recorded crosstalk, whereas a resultant smaller crosstalk light pattern means a counteraction by the compensating signal. In this way various bands of single channel modulations may be recorded, with the crosstalk known both in magnitude and phase angle (the latter only as far as "in phase" or "out of phase" with the main signal is concerned). Tracing

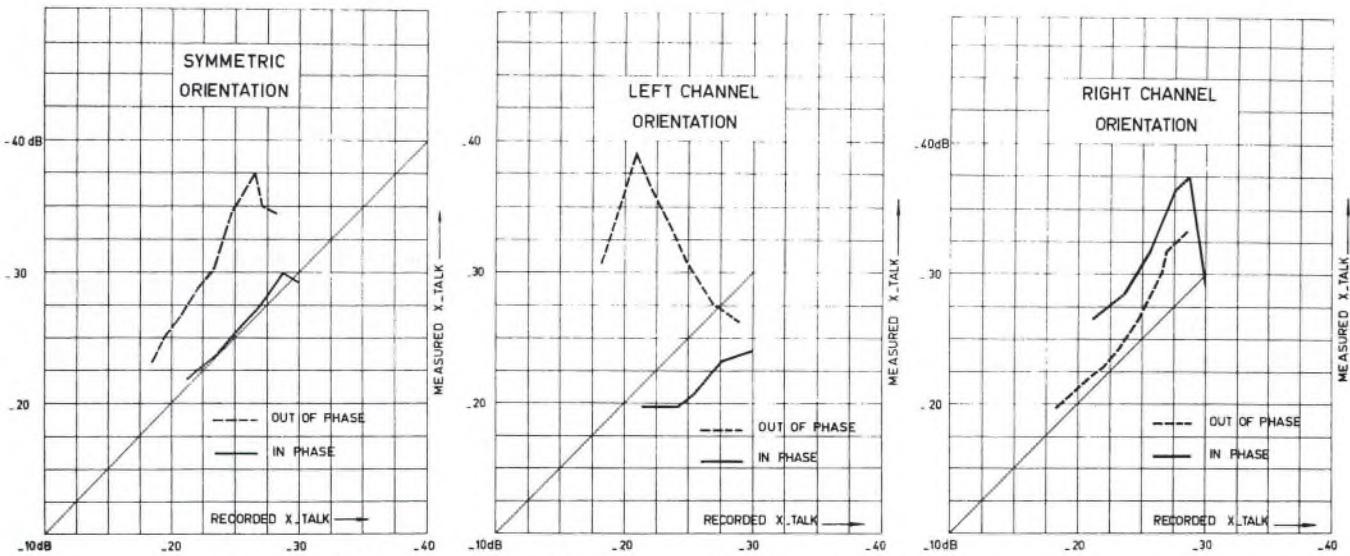


FIG. 17. Erratic behavior of stereo pickup on crosstalk of different magnitude and phase.

such a special test record with a stereo pickup leads to very erratic measured results. It may occur that the test band having the lowest level of recorded crosstalk produces the highest crosstalk in the pickup output and vice versa. Figures 17a, b and c show the behavior of the pickup used in this test as a function of the recorded crosstalk. The pickup used had an axes orientation with an included angle of 85° (the method for determining this will be dealt with later on). Figure 17a shows the crosstalk behavior for a symmetric alignment of the pickup head; Figure 17b for exact 45° orientation of the left channel system; figure 17c for the right channel system. The measuring results of all bands where the recorded crosstalk is in phase with the main signal have been joined into a curve separate from one other curve for the out-of-phase recorded crosstalk levels. If the pickup did not contribute to crosstalk at all, the measured values would equal the recorded ones (the straight line in the three figures). The curves do not show any tendency towards a fundamental behavior; they only prove that any statement about the separation capacity of a stereo pickup is doubtful when the separation figure has been obtained by utilizing a test record with normal single-channel modulation.

III.2 New method of measuring separation

When a lateral signal is traced by one system of a stereo pickup, the greater part of this system's output is caused by the decomposed component E_h of the lateral signal in the direction of this system's greatest sensitivity. Part of the output signal, however, is also caused by crosstalk e_h coming from the movements of the opposite system. A similar thing occurs when the pickup traces a pure vertical signal (E_v and e_v). In this case, however, the phase of the cross-talk voltage is reversed. If the lateral signal is recorded with the same velocity as the vertical one, $E_h = E_v$ and $e_h = -e_v$. In one case, the system's output is represented by

$$E_1 = E_h + e_h$$

in the other by

$$E_2 = E_v - e_v.$$

We may therefore write:

$$\frac{E_2}{E_1} = \frac{(E_v - e_v)/(E_h + e_h)}{[1 - (e_v/E_v)]/[1 + (e_h/E_h)]} = \frac{(E_v - e_v)/(E_h + e_h)}{[1 - (1/S)]/[1 + (1/S)]} \quad (9)$$

The separation is defined by the ratio $S = E_h/e_h = E_v/e_v$. Thus,

$$M = E_2/E_1 = [1 - (1/S)]/[1 + (1/S)] \quad (10)$$

or the separation is:

$$S = (1 + M)/(1 - M). \quad (11)$$

This relation is shown in graphical form in Fig. 18. For $M = 1$, infinite separation is obtained. The procedure is then to measure the output voltage of a pickup system using a pure lateral signal and compare it with the output obtained from a pure vertical signal of exactly equal magnitude. A difference of, for instance, 1 db means a separation figure of 25 db for the measured system as given by Fig. 18. This value of separation is now obtained with a method which avoids the troubles inherent in one-wall modulations. The creation of a really one-wall modulation, i.e., without any signal whatsoever in the opposite wall, will probably remain an utopian ideal. Pure lateral and vertical signals, however, are practically possible to cut. A further advantage of this method is that the two voltages measured per system are of the same order of magnitude. Phase angles have purposefully been neglected in the derivation of the above formulae. It remains to be seen whether the theory still holds for frequencies near possible pickup resonances. Between 100 cps and 5000 cps the practical results do agree with the theoretical expectations, as will be shown in the following paragraph. Care should be exercised with high frequencies due to the pinch effect in lateral recording.

III.3 Measuring the orientation of pickup sensing axes

It may be welcome to have a method for determination of

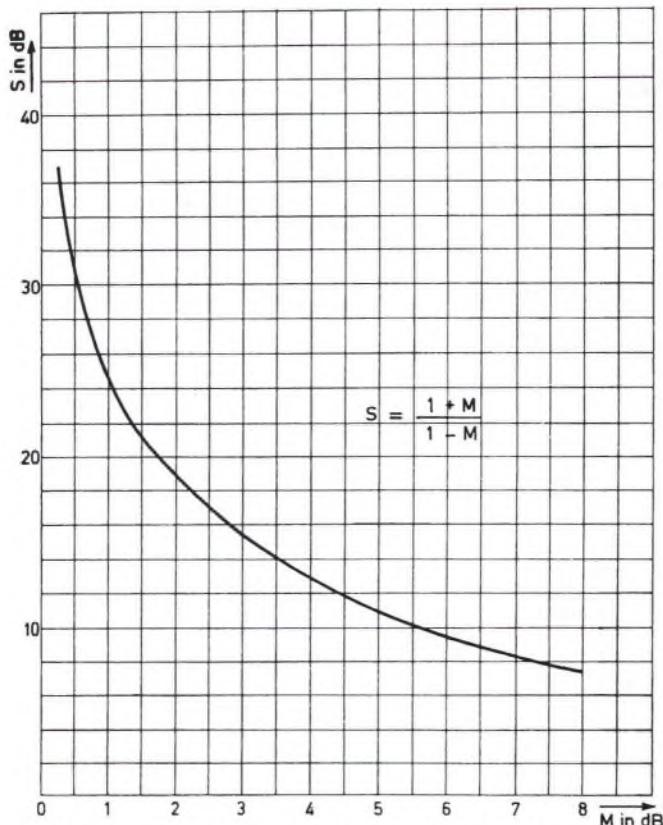


FIG. 18. Graph for calculation of pickup separation by means of a new method of measurement omitting *recorded* crosstalk level and phase.

the orientation of axes of a stereo pickup. In Fig. 19 the orientation of axes is represented by the included angles δ

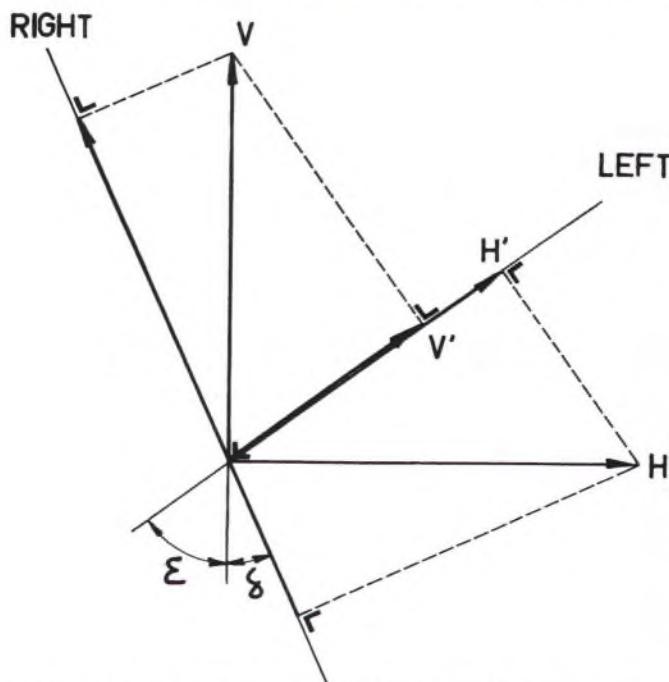


FIG. 19. Vector presentation of a measuring method for orientation of pickup axes.

and ϵ relative to the normal. Tracing a lateral signal, represented by vector H , with one system, for instance the left-hand system, produces an output created by the vector of movement H' . Likewise, tracing a vertical signal (vector V of equal magnitude as H) produces V' . It can be shown (see Appendix D) that

$$\tan \epsilon = H'/V'. \quad (12)$$

Consequently, from the ratio M discussed in the former paragraph, the orientation angles δ and ϵ may be calculated.

$$\begin{aligned} \text{If } & H' = V' \quad \epsilon = 45^\circ \\ & H' > V' \quad \epsilon > 45^\circ \\ & H' < V' \quad \epsilon < 45^\circ. \end{aligned}$$

In Fig. 20 the angular difference $\Delta\epsilon$ from 45° is plotted as a function of measured ratio $M = H'/V'$.

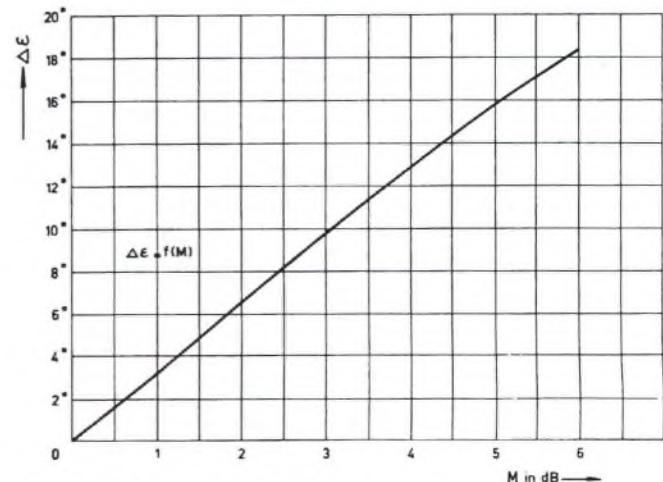


FIG. 20. Graph for calculation of error in orientation of pickup axes.

It may be remarked that the actually measured system outputs comprise both main signal and crosstalk, on which the theory of the preceding paragraph has been based. The crosstalk has been neglected in the derivation of Eq. (12), but it can be shown that the error caused by this negligence is of no importance.

III.4 Practical results

Rotating a pickup head along its longitudinal axis and measuring ratio M , the separation of the pickup head may be determined as a function of the rotative variation in degrees. This has been done for a certain type of pickup using special test records with carefully cut bands of lateral and vertical signals (equal velocities) at frequencies of 100, 1000, 2340 and 5200 cps.

From the ratio M found with zero alignment of the pickup head, the orientation of axes has also been determined. As was to be expected, the orientation of axes *differs for each frequency* (see Fig. 21). For each frequency, the agreement with theoretical expectations is very satisfactory. Take the case of 1000 cps, for instance. The calculated orientation is $42^\circ 42'$ for the left-hand channel and $41^\circ 43'$

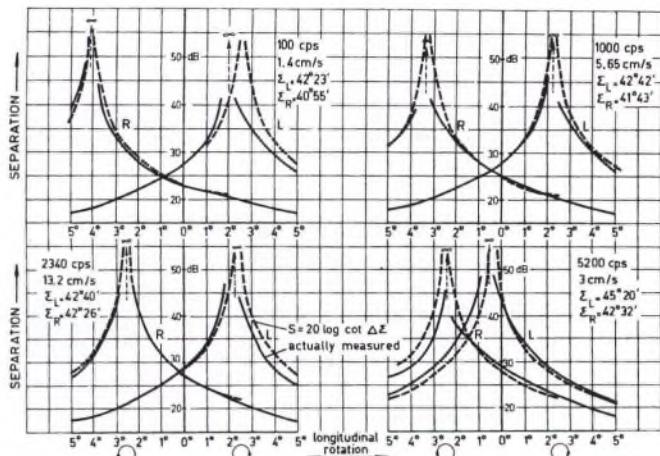


FIG. 21. Actually measured variation of separation at four frequencies due to non-coincident orientation of axes, compared with theoretical expectations.

for the right-hand channel relative to the normal. This means that the pickup head has to be rotated about 2.3° clockwise for optimal separation in the left channel and about 3.3° counterclockwise for the opposite system. Note that the rotation curves do indicate these values and directions within an error of less than 0.5° .

$$\Delta = \text{theoretical separation} \quad (S = 20 \log \frac{1+M}{1-M})$$

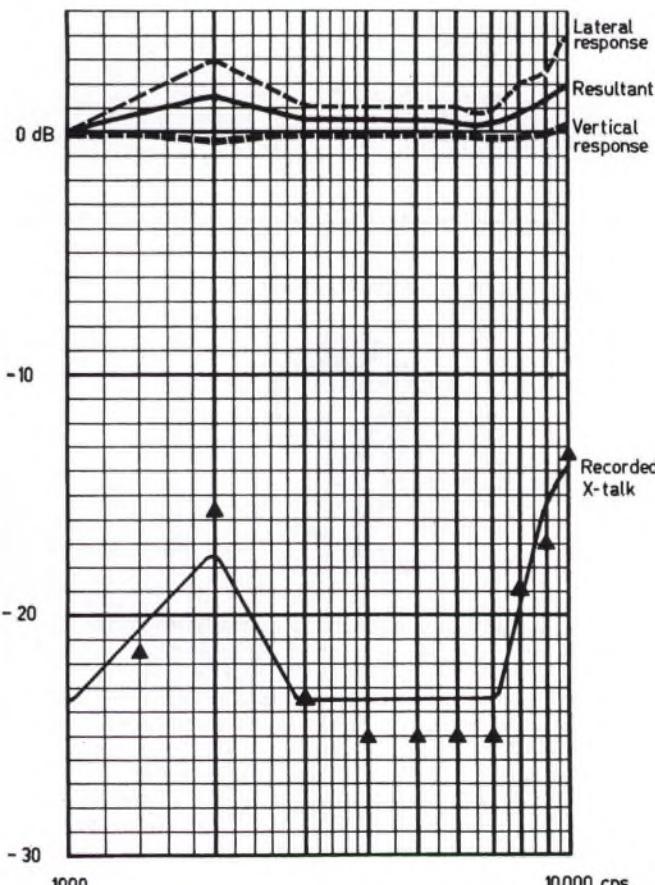


FIG. 22. Increase of recorded crosstalk in matrixed 45/45 cutting due to differences in response of either channel of a 0/90 cutter.

IV. DIRECT 45/45 RECORDING VERSUS MATRIXED 0/90 CUTTING

A 45/45 recording made by means of a 0/90 cutter to which matrixed signals, to wit $(L + R)$ and $(L - R)$, are fed, may show up rather bad crosstalk rejection rates if the lateral and vertical channels of the 0/90 cutter head differ in frequency response. A perfect 45° recording is only possible if both channels (H and V) are exactly identical in behavior. This is in fact a reverse application of the theories of Chapter III.

From Eq. (11):

$$S = (1 + M)/(1 - M)$$

where S stands for separation or crosstalk rejection and M is the ratio of lateral to vertical signal amplitude, the recorded crosstalk resulting from 45° recording with 0/90 cutters may be calculated. If for a certain frequency the response in lateral and vertical channels differs 3 db, the 45° recording at this frequency will show a theoretical crosstalk value of $15\frac{1}{2}$ db! Since a frequency response deviation of \pm or -1 db is normally allowable, the lateral and vertical excursions of a 0/90 cutter may differ 2 db, meaning a "one-wall" recording having about 19 db crosstalk.

The trueness of this theory is proved by the following test. With a 0/90 cutter a test record has been cut. The test was preceded by a frequency response check of both lateral and vertical channels by means of the light pattern method.² The lateral channel was far from flat, having a maximum deviation of $+4$ db relative to 1000 cps. The vertical channel proved to be substantially flat. The resultant 45° recording of matrixed responses was checked, again with light patterns, and showed to have a response flat within 2 db. The recorded crosstalk was also measured with the light pattern method. The measured results are shown in Fig. 22 together with the predicted crosstalk levels as determined from the abovementioned equation. The agreement with the actually recorded crosstalk levels is striking.

It should be borne in mind that this phenomenon applies to all matrixed orthogonal cutting systems, i.e., differences in frequency response of either channel of a 45/45-cutter show up as deviations from true lateral and vertical recordings by feeding in matrixed signals. But since we are concerned with the standardized way of stereo recording following the 45/45 system, it is plainly clear that the performance of a 0/90 cutter must be very good indeed, particularly as far as equality in frequency responses is concerned. A direct 45/45 recording comes out favorably because dif-

APPENDICES

A. Equations (1) and (2) in Paragraph I.2.1

Refer to Fig. 1.

$$\begin{aligned} \text{in } \Delta TCD: \cot \epsilon &= TC/CD \\ \text{in } \Delta TEA: \cos \theta_e &= TE/TA \\ \text{in } \Delta TEC: \cos \theta_p &= TE/TC \end{aligned}$$

² C. R. Bastiaans and J. v. d. Steen, "Optical measurement of recorded velocities on stereo disk records,"—Philips Technical Review, 23, nr. 3, 89 (1961-62).

ferences in channel responses only show up as such, and possible deviations from true lateral or vertical recordings with such a cutter are mostly of minor concern.

Assuming $\beta = 45^\circ$ (true orientation of cutter axes):

$$CD = AB = TA$$

so that

$$\cot \epsilon = TC/TA = \cos \theta_c / \cos \theta_p. \quad (\text{Eq. 1})$$

Specific case for $\theta_c = 0^\circ$:

$$\tan \epsilon = \cos \theta_p. \quad (\text{Eq. 2})$$

B. Equations (3), (4), (5) and (6) in Paragraph I.2.3

Refer to Fig. 23.

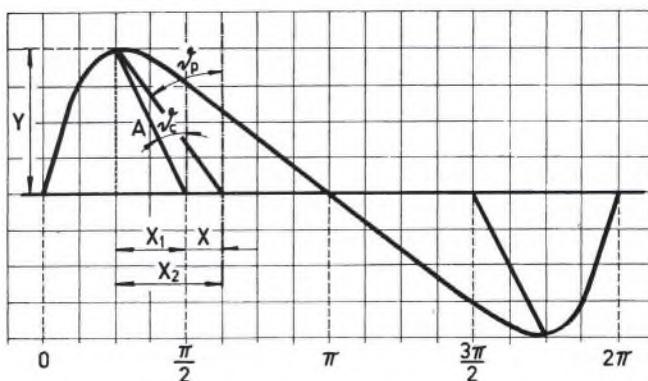


FIG. 23. Tracking distortion caused by a difference between vertical cutting and playback angles.

Cutting a sinewave with a vertical angle of θ_c introduces a distortion, the maximum amplitude of which is represented by the relation between the horizontal shift x_1 and the vertical projection y of the signal amplitude A .

Tracing the distorted sinewave along a vertical angle θ_p equal to θ_c eliminates the distortion. Should the playback angle θ_p differ from θ_c , the difference x between the horizontal projections x_1 (for cutting) and x_2 (for playback) means a relative horizontal shift, i.e., *tracking* distortion is introduced. This distortion may be expressed as follows:

$$\% D_2 = (2\pi x/\lambda) \cdot 100\%$$

where D_2 = amplitude of second-order distortion product and λ = the wavelength of the recorded signal.

$$\begin{aligned} x &= x_2 - x_1 \\ x_1 &= A \sin \theta_c \\ x_2 &= y \tan \theta_p \\ y &= A \cos \theta_c \end{aligned} \quad \rightarrow x_2 = A \cos \theta_c \tan \theta_p$$

so that we may write:

$$\begin{aligned} x &= A (\cos \theta_c \tan \theta_p - \sin \theta_c) = Aa \\ 2\pi x/\lambda &= (2\pi f Aa)/V = (v/V)a \end{aligned}$$

Thus:

$$\% D_2 = v/V (\cos \theta_c \tan \theta_p - \sin \theta_c) \times 100\% \quad (\text{Eq. 3})$$

If $\theta_p < \theta_c$, $x_2 < x_1$ so that $x = x_1 - x_2$ and:

$$\% D_2 = v/V (\sin \theta_c - \cos \theta_c \tan \theta_p) \times 100\% \quad (\text{Eq. 4}).$$

Specific cases for $\theta_c = 0^\circ$ or $\theta_p = 0^\circ$ give Eqs. (5) and (6) respectively.

C. Equation (8) in Paragraph II.1

Refer to Fig. 11.

Maximum allowable stylus velocity is

$$v_{max} = V \tan (90^\circ - \phi - \theta_c) = V \cot (\phi + \theta_c) \quad (\text{Eq. 8});$$

for $\theta_c = 0^\circ$ this reduces to

$$v_{max} = V \cot \phi.$$

The decrease in maximum allowable velocity due to an artificial cutting angle θ_c is therefore

$$V \cot (\phi + \theta_c) / V \cot \phi.$$

For $\phi = 45^\circ$ this equation reduces to the equation used in Fig. 12,

$$\tan (45^\circ - \theta_c).$$

D. Equation (12) in Paragraph III.3

Refer to Fig. 19:

$$\begin{aligned} \cos \epsilon &= V'/V \\ \sin \epsilon &= \cos (90^\circ - \epsilon) = H'/H. \end{aligned}$$

Since $H = V$, $\tan \epsilon = H'/V'$ (Eq. 12).

THE AUTHOR



C. R. Bastiaans was born in the former Netherlands East Indies in 1924. He attended school in Holland and received the diplomas of radio engineer and electrical engineer. In 1950, after five years of officership in the army, he joined the Philips telecommunication department as a technical-commercial engineer. In 1957, he became head of the electro-mechanical development group of Philips Phonographic Industries.

Applications for several of Mr. Bastiaans' patents in the audio field have been made. He is a contributor to the Philips Technical Review and has been published previously in the Audio Engineering Society Journal.

ON GEOMETRIC CONDITIONS IN THE CUTTING AND PLAYING OF STEREO DISCS

DUANE H. COOPER

University of Illinois, Urbana, Illinois

With regard to C. R. Bastiaans' paper, "Further Thoughts on Geometric Conditions in the Cutting and Playing of Stereo Discs" (*J. Audio Eng. Soc.* 11, 6, Jan. 1963)* I think it is remarkable that Mr. Bastiaans should present measured harmonic distortion resulting from vertical tracking angle error running to dozens of percentage points, and, in the same paper, claim to have "established that the effects of vertical tracking error are not as serious as is sometimes assumed." If he had been aware of the intermodulation distortion, first analyzed in the pioneering work of Baerwald,¹ and the crossmodulation that arises in the same manner for the stereo case,² as a result of tracking angle error, he might have been more easily shocked. Such distortion is frequency-dependent, runs to much greater percentages and is audibly distressing.

Mr. Bastiaans is impressed with cutting difficulties, but no such difficulties are intrinsic. It is not necessary that the cutting stylus have non-standard angles between its facets to cut with a vertical angle different from zero degrees. It is only necessary that the stylus be properly mounted. In the Westrex 3C cutter, a 23 degree angle is used, as everybody knows,³ yet the standard-stylus facets bear the same angles as always with respect to the vertical and to the groove axis. There is no real question of inherent vs artificial "slant," here. It is possible, of course, that a redesigned stylus mount may be required if the vertical angle of the cutter is to be changed, but in the case of the Westrex 3C a more radical redesign would be needed anyway if a 15 degree angle were to become the standard.

Part of Mr. Bastiaans' fears about cutting difficulties concern the tendency of the stylus to "dig" into the lacquer. He should have noted that this tendency depends as much upon the angle that the line from stylus tip to stylus pivot makes with the groove axis, as it does upon the angle from cutting facet to groove axis.

REFERENCES

1. H. G. Baerwald, *J. Soc. Mot. Pict. Eng.*, 37, 591 (1941).
2. D. H. Cooper, "Tracking Distortion as Phase Modulation," *IRE Trans. on Audio* (to be published).
3. C. C. Davis and J. G. Frayne, *Proc. IRE*, 46, 1686 (1958).

ON GEOMETRIC CONDITIONS IN THE CUTTING AND PLAYING OF STEREO DISCS

J. V. D. STEEN

Philips' Phonographic Industries, Baarn, Netherlands

As Mr. Bastiaans has left our company, it is my task to answer Mr. Cooper's letter (*JAES* 11, 191, Apr. 1963) concerning the article: "Further Thoughts on Geometric Conditions in the Cutting and Playing of Stereo Discs" (*JAES* 11, 6, Jan. 1963).

It is a well-known fact that tracking error causes non-harmonic frequency-dependent automodulation-distortion as well. This subject, however, needs a special study. In the article of Mr. Bastiaans only the harmonic distortion due to vertical tracking error has been taken into account, and we are very interested in the results of Mr. Cooper's investigations of tracking distortion as phase modulation.

When Mr. Bastiaans claims (p. 9, point 1.4) that "the effects of vertical tracking error are not as serious as is sometimes assumed," though tracking distortion can reach considerably high values, he bases this remark on the fact that *tracking distortion*, which occurs at the same time during playback, *dominates the effect of tracking distortion*, as mentioned on p. 7 (point 1.2.3) and p. 9 (righthand column, second paragraph). For higher frequencies tracking distortion increases, as this type of distortion is frequency-dependent, while tracking distortion is only dependent on cutting velocity, groove velocity and tracking error (p. 7, point 1.2.3, giving formulae identical to the formulae of Baerwald and Bauer for harmonic distortion). Thus, tracking distortion is only of minor importance in comparison with tracking distortion occurring at the same time.

FURTHER COMMENTS ON GEOMETRIC CONDITIONS IN THE CUTTING OF STEREO RECORDS

DUANE H. COOPER

University of Illinois, Urbana, Illinois

Mr. v. d. Steen is quite correct in supporting Mr. Bastiaans in maintaining that tracing distortion dominates tracking distortion at the higher frequencies, if only harmonic distortion be considered. Sticking by his own "ground rules," however, and considering only harmonic distortion it can be shown that tracking distortion dominates that part of the spectrum responsible for the impression of overall loudness, or level, of the separation (vertical) signal. Thus, placing a bound on the combined distortion requires a sharp reduction in separation, or stereo effect, if the tracking error itself is not to be removed, as by standardizing actual (not fictive) cutter angles.

It is easy to see that tracing distortion affects the recorded waveform $f(x)$ in such a way that, if the waveform obeys $f(x) = f(-x)$, then so does the distorted waveform $f'(x)$, i.e., one has $f'(x) = f'(-x)$, as well. Thus, if the test waveform were a cosine, the distorted waveform may be represented in a cosine series. For small amounts of distortion, the dominant second harmonic term is of relative magnitude¹ $a/4\rho_{min}$, in which ρ_{min} is the radius of curvature at the places of sharpest curvature in the waveform, and a is the radius of the reproducing stylus tip. However, the application of tracking distortion, with its characteristic phase automodulation,² generates a dominant sinusoidal second harmonic term, when applied to the cosine, of magnitude $(v/v_g) \tan \phi$, if a truly vertical cut had been made. Here v/v_g is the peak slope of the waveform, v_g being the groove speed, v the peak recording velocity, and ϕ is the tracking error. Thus, the two second harmonic terms are in phase quadrature and their energies, not their amplitudes, add.

These results can be combined to predict the total distortion due to both effects at various frequencies. We should confine ourselves to small distortion levels, so that we need take into account only these kinematic effects. At great distortion levels, dynamic effects come into play.³ The "scare value" of large distortion figures has probably been adequately exploited by now, and, in any case, no one wants to listen to large amounts of distortion.

Five percent is often considered satisfactory, although no one in the high fidelity electronic components industry, probably, is satisfied with more than one or two percent. However, if we can provisionally settle on a figure like five percent, we can draw up a velocity de-rating curve, as in Fig. 1, showing the maximum peak

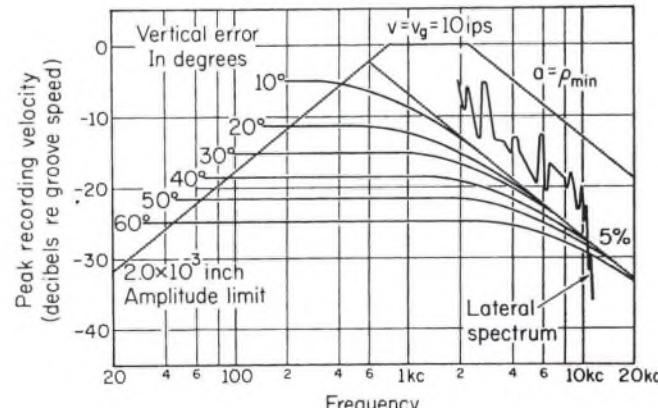


FIG. 1. Peak recording velocity vs frequency.

1. M. S. Corrington, "Tracing Distortion in Phonograph Records," *RCA Rev.* 10, 241 (1949).

2. D. H. Cooper, "Tracking Distortion as Phase Modulation," *IEEE Trans. on Audio* (in press).

3. J. Walton, "Stylus Mass and Distortion," *Wireless World* 69, 171 (1963).

velocities that can be recorded, for the vertical, if five percent is not to be exceeded.

The 2-mil amplitude limit is somewhat arbitrary and has nothing to do with the five percent figure. The same is true of the $v = v_g$ limit and the $\rho_{min} = a$ limit. The 60-degree contour is interesting in that its flat portion is very nearly at the rms level of one centimeter per second.

The five percent limits for lateral recording are much more generous, almost extending to the $\rho_{min} = a$ contour. That 6-db-per-octave slope is not terribly confining, since the spectra of most program material will conform, even after the RIAA equalization. The portion labeled "lateral spectrum" was taken from data published by Walton;⁴ levels corresponding to loudness peaks are not shown, since several minutes of data are represented in an average sense.

A 10-db downward shift will make the contours applicable to a 1.6 percent distortion figure. Sliding the contours to the right or left will allow interpretations for other groove speeds. Replacing $\tan \phi$ by $\sin \phi \cos (\phi + \psi)$ will allow calculations for cutter angles, ψ , other than zero degrees, if ϕ is maintained as the tracking error.

If recording engineers are willing to so sharply restrain the level of the vertical cut, as shown, then tracking error, combined with tracing error, can be "lived with." Rough measurements on some of the better discs show such small vertical levels, and the separation is still pleasingly noticeable. It is submitted, however, that the 30 degree contour, which may be typical for European pressings played with a 15-degree cartridge, imposes a severe limitation over a great and important part of the audio spectrum.

4. J. Walton, "Versatile Stereophonic Pickup," *Wireless World* 67, 407 (1961).

A Study of Program-Level Overloading in Phonograph Recording*

J. G. WOODWARD AND E. C. FOX

RCA Laboratories, Princeton, New Jersey

Displacement, slope, and curvature overload limits in phonograph recording are defined in terms of groove and stylus dimensions which are currently standard for 45/45 stereodisks. Instrumentation is described for making independent measurements of the occurrence of each type of overload. Data are presented showing the number of occurrences of each type of overload as a function of program level for several musical selections.

1. THE NATURE OF OVERLOADING

IN every system for the transmission or storage of signals serious distortion of the signal occurs whenever the signal level exceeds certain limits which are characteristic of the system. In phonograph recording, level limits are determined by purely geometrical factors involving the dimensions of the record groove and of the recording and playback styli. Three types of overload, namely, displacement, slope and curvature, exist. These are well known and have been described in the technical literature, both for lateral recording¹ and for 45/45 stereophonic recording.^{2,3} However, they will be reviewed in this section in order to relate them quantitatively to data to be presented in subsequent sections, where measurements of the frequency of occurrence of overloads and their dependence on level will be given for recorded music.

Displacement Overload

Displacement overload occurs when the lateral component of the peak displacement of the recording stylus is sufficient

to cause adjacent grooves to partially overlap. The basis for determining the displacement-overload limit in 45/45 stereo recording is illustrated in Fig. 1, which portrays a

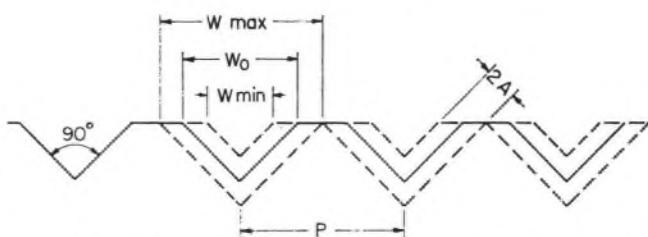


FIG. 1. Vertical section through several 45/45 stereo grooves, showing limits of displacement.

section through several adjacent grooves. The center-to-center separation of unmodulated grooves is P , with $1/P$ being the number of grooves per inch. The minimum allowable groove width is W_{min} , the unmodulated groove width is W_o , and the displacement amplitude along each 45° direction is A . The maximum groove width, W_{max} , is taken as that for which adjacent grooves of this width just touch at the record surface. Therefore,

$$W_{max} = P$$

When the groove walls have a 90° included angle the displacement limitation is given by

* Presented October 16, 1962 at the Fourteenth Annual Fall Convention of the Audio Engineering Society, New York.

¹ W. S. Bachman, *Electronic Ind.* 4, 86 (July 1945).

² J. W. Stafford, *J. Audio Eng. Soc.* 8, 162 (July 1960).

³ Erling P. Skov, *J. Audio Eng. Soc.* 8, 12 (Jan. 1960).

$$A_{max} = (1/4\sqrt{2})(P - W_{min}).$$

Taking $W_{min} = 0.001$ in., the maximum displacement in inches along a 45° direction which will not cause overcutting of grooves is, therefore, in terms of the groove pitch

$$A_{max} = .177(P - 10^{-3}). \quad (1)$$

For sinusoidal modulation the displacement overload may be expressed in terms of stylus velocity by the relation

$$v_{max} = 2\pi f(2.54)A_{max} = 2.83 f(P - 10^{-3}) \quad (2)$$

where v_{max} is the velocity amplitude in centimeters per second corresponding to displacement overload, and f is the frequency of the signal in cycles per second. Equation (2) is plotted for various values of $1/P$ in Fig. 4. In practice some overcutting may be accepted as long as the groove depth is not reduced to a value which will no longer reliably contain the pickup stylus or which will cause adjacent-groove crosstalk. The amount of overcutting which may be tolerated depends on the number of grooves per inch and the groove depth. It seems wisest to reserve this overcutting as a margin of safety rather than to include it in our definition of the displacement overload limitation.

Slope Overload

Slope overloading occurs when the program level and waveform are such that the slope of the modulated groove wall becomes greater than the slope of the trailing faces of the recording stylus. This is illustrated in Fig. 2 for the

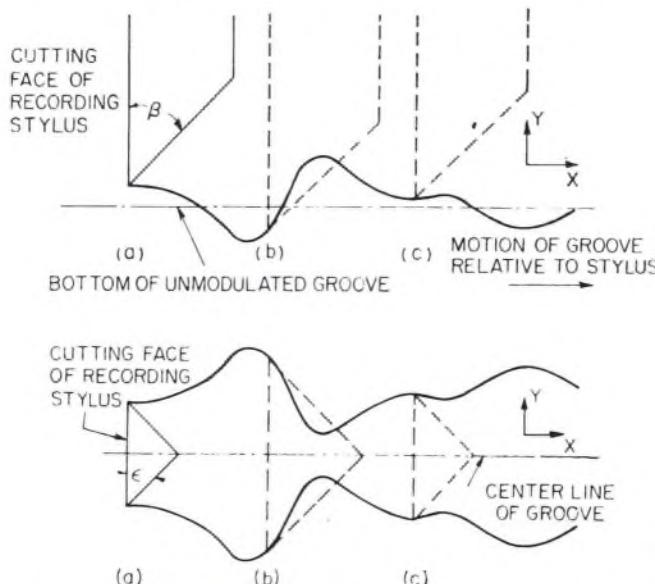


FIG. 2. Sections through a vertically-modulated groove, showing a condition for slope overloading. *Top:* vertical section through the center of the groove. *Bottom:* horizontal section at the surface of the record.

case of a vertically-modulated groove. The upper part of the figure depicts a vertical section of the groove and the recording stylus in the central plane of the groove. The lower portion shows a horizontal section of the stylus and

groove in the plane of the record surface. The stylus is indicated at three locations. It is clear that a stylus of the shape shown in Fig. 2 cannot possibly record the modulation indicated at location (b), and a distorted signal must result. Similar considerations apply to a laterally-modulated groove and to either of the 45/45 stereo modulations taken individually. In general terms, if y is the stylus (or groove) displacement, the slope of the groove modulation is dy/dx , where x is measured in the direction of groove motion relative to the stylus. If slope overloading is to be avoided, the groove-modulation slope in the vertical plane must meet the condition

$$(dy/dx) \leq \operatorname{ctn} \beta$$

where β is the back angle of the recording stylus, and y is now taken as vertical displacement. Similarly, in the horizontal plane

$$dy/dx \leq \operatorname{ctn} \epsilon$$

where ϵ is the edge angle of the recording stylus and y is now taken as the lateral displacement.

With recording styli of currently-standard dimensions, both β and ϵ are nearly equal to 45° . Hence, the slope-overload limit is given by

$$dy/dx = 1.$$

If z is the displacement along one of the 45/45 directions, $y = \sqrt{2}z$ and

$$dz/dx = 1/\sqrt{2}. \quad (3)$$

The slope-overload limit may be expressed in terms of the instantaneous stylus velocity, v , and the linear groove speed, V_G , as

$$v_{max} = V_G/\sqrt{2},$$

and in terms of the groove diameter, D , and the rotational speed of the record, Ω , as

$$v_{max} = 2.54\pi D\Omega/60\sqrt{2} = 0.0942D\Omega. \quad (4)$$

In Eq. (4), v_{max} is the maximum permissible velocity in centimeters per second in a single 45/45 channel when D is given in inches and Ω in rpm. Equation (4) is plotted in Fig. 4 for the innermost and outermost grooves on $33\frac{1}{3}$ rpm, 12-in. records.

Stereo recorders are usually constructed in such a way that the pivot point of the recording-stylus holder does not lie in the plane of the record surface and, when driven for vertical modulation, the stylus tip moves along a line making an angle with the true vertical. The effect of this "tilt angle" in giving rise to distorted vertical modulation in the recorded groove has been described by Davis and Frayne⁴ and Skov,⁵ and will not be discussed further here. The effective modulation of the relative groove-stylus speed due to the tilt angle is in a sense which reduces slope overloading for vertical modulation. However, lateral modulation is not

⁴ C. C. Davis and J. G. Frayne, *Proc. IRE* 46, 1686 (1958).

⁵ Erling, P. Skov, *loc. cit.*

affected, and Eq. (4) still must be taken as the slope-overload limit.

Curvature Overload

Curvature overloading occurs when the radius of curvature of the groove modulation is less than the tip radius of the pickup stylus. When the overload limit is exceeded, the pickup stylus may be subjected to mechanical impacts with the groove walls which result in severe non-linear distortion and shock excitation of the mechanical system of the pickup. Curvature overloading is illustrated in Fig. 3. The

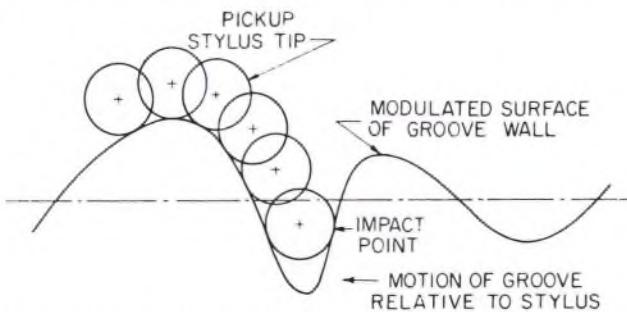


FIG. 3. Section of groove in plane of modulation, illustrating a curvature-overload condition.

cross section of the tip of the pickup stylus, represented by a sphere, is shown in several locations on the modulated surface of a groove wall. The lowest stylus location in the diagram depicts a severe case of curvature overloading.

Whereas the displacement- and slope-overload limits represent a rather definite point at which overcutting and distortion occur, respectively, the curvature-overload limit described here represents only an arbitrary upper limit of tracing distortion. For some types of program material the tracing distortion will have become intolerable even before the curvature-overload point is reached. However, the curvature-overload concept is useful in analytical work because of its mathematical convenience and definiteness.

If x is distance measured longitudinally along the recorded groove and y is the displacement taken either in a vertical direction or along one of the 45/45 stereo axes, the radius of curvature, ρ , of the groove modulation is

$$\rho = [1 + (dy/dx)^2]^{3/2} / (d^2y/dx^2). \quad (5)$$

If r is the tip radius of the pickup stylus, the curvature-overload limitation occurs when $\rho = r$. For a sinusoidal modulation the radius of curvature should not be less than the minimum value of

$$\rho_{min} = (\lambda/2\pi)^2 \cdot (1/A)$$

where λ is the wavelength of the signal recorded in the groove and A is the displacement amplitude. The curvature-overload limit in this case may be expressed as

$$A_{max} = (\lambda/2\pi)^2 \cdot (1/r), \quad (6)$$

where A_{max} is the displacement amplitude for which the radius of curvature of the modulation is just equal to the

stylus tip radius. When expressed in terms of peak stylus velocity, the curvature overload limitation is given by

$$v_{max} = 2\pi f A_{max} = V_g^2 / 2\pi f r,$$

whence

$$v_{max} = 2.54 / 2\pi f r \cdot (\pi D \Omega / 60)^2 = (1.10 \times 10^{-3} D^2 \Omega^2) / f r. \quad (7)$$

The symbols D and Ω were defined earlier. Dimensions of D and r are in inches, while v_{max} is in centimeters per second. Equation (7) is plotted in Fig. 4 for the innermost and outermost diameters of a 33 1/3, 12-in. record, assuming a pickup stylus-tip radius of 0.0007 in.

It should be remembered that Eq. (7) is applicable only at the peak of a sinusoidal modulation. We must return to Eq. (5) for an exact expression of curvature overload applicable to any point of any modulation waveform. The space coordinate, x , may be replaced by the time coordinate to yield a curvature-overload expression in terms of stylus velocity and acceleration as follows:

$$\rho = [1 + (V_a \cdot dy/dt)^2]^{3/2} / (V_g^2 \cdot d^2y/dt^2) = r. \quad (8)$$

In general the relation between velocity, dy/dt , and acceleration, d^2y/dt^2 , is not known. Measurements of curvature overloading for musical program have been made using both the approximate expression, Eq. (7), and the exact expression, Eq. (8), as the basis. These measurements will be discussed in a subsequent section.

Discussion

The overload limits for the three types of overload are plotted in Fig. 4 for sinusoidal signals. Also shown is the

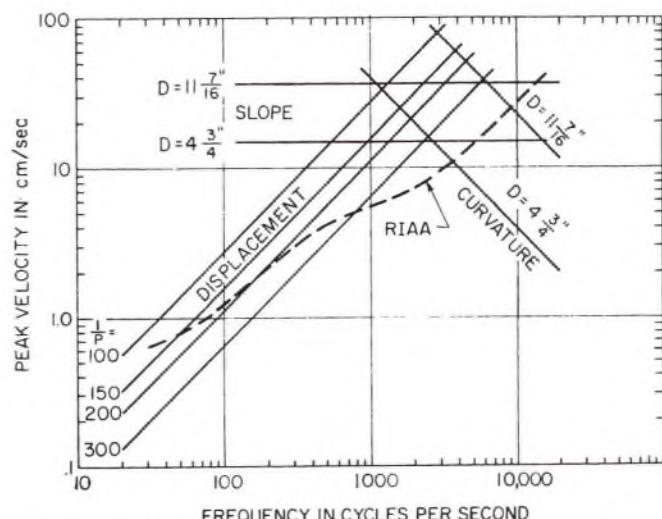


FIG. 4. Overload limits for displacement, slope and curvature for 33 1/3 rpm rotational speed and for a 0.7-mil pickup stylus tip radius. D = groove diameter; $1/P$ in grooves per inch.

recording stylus velocity as a function of frequency when the recording level has a peak velocity of 5.5 cps at 1000 cps in a single stereo channel, and the RIAA recording equalization is used. It is evident from the analysis that

even for this standard recording level, curvature overloading will occur frequently unless the power of the high-frequency components in the program is low relative to the mid-band components. Assuming that program peaks are no more than 10 db higher than the standard level as plotted, slope overloading should occur only rarely. Displacement overloading should never occur if the pitch is properly controlled.

In the sections which follow, an experimental set-up will be described which was used to count the occurrence of each of the three forms of overload for several types of musical program. Curvature overloading occurs more frequently and is more complex analytically than the other types of overload. It will be treated first.

2. MEASUREMENT OF OCCURRENCES OF CURVATURE OVERLOAD

Instrumentation

A system terminating in an electronic counter was set up to count the number of occurrences of acceleration peaks in several musical passages which would cause curvature overloading to occur on a phonograph record for specified recording conditions. The musical passages were selected portions of two-channel stereotapes. The selections, described in Table I, were played at 7½ ips on a tape recorder which

TABLE I. Musical selections used for the measurement of occurrences of curvature overload.

Selection No.	Description	Duration (seconds)
I	Symphony orchestra playing a Strauss waltz; high-level violin passage.	60
II	Paul Lavalle's concert band; passage contains muted trumpets which are responsible for most of the curvature overloads.	190
III	Paul Lavalle's concert band; no muted trumpets.	184
IV	Bagpipes and drums; high level.	97
V	Concert orchestra; light-opera overture.	122
VI	Female "pops" vocalist; low-level orchestral accompaniment; consonants accentuated.	191
VII	Operatic soprano soloist; symphony orchestra accompaniment.	74
VIII	Symphony orchestra; full orchestra; heavy percussion.	87
IX	First portion, only, of Selection VI.	60
X	Sauter-Finnegan Orchestra	60
XI	Pipe organ; strong pedal stops	53

had been adjusted to give a flat response-frequency characteristic from 30 to 15,000 cps when playing the RCA Victor Standard Test Tape No. 12-5-61T. The music outputs from the two stereo channels were added to give the entire program content in a single channel. This combined signal was then processed and fed to the counter. A block diagram of the basic system used in measuring curvature overloads is shown in Fig. 5.

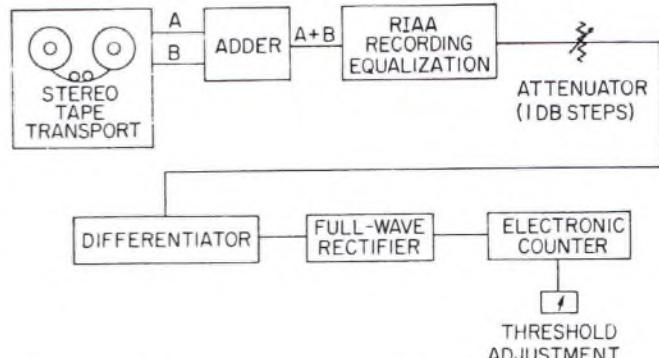


FIG. 5. Block diagram of system for counting curvature overloads.

The first step in processing the program was to subject it to the RIAA recording-equalization characteristic which is shown in Fig. 4. This equalized voltage would be proportional to the velocity of the recording stylus in a disk-recording system. When measuring curvature overloading the equalized signal was differentiated to give a voltage which would be proportional to the acceleration of the recording stylus. The equalized and differentiated signal was fed to a full-wave rectifier, and the rectified waveform went to the counter. In a vertically-modulated groove or in a single 45/45 stereo channel, curvature overloading occurs for only one polarity of the signal. However, for the present tests, the full-wave rectification gave overload counts for both polarities. Hence, statistically, the number of occurrences of overloading should actually be only half the total number counted. It was necessary to count only those program peaks which exceeded some limit which was determined by the type of overload being studied and the recording conditions specified. This meant setting a threshold at the counter corresponding to the overload limit. In the case of curvature-overload measurements the threshold calibration was made in the following way. With the step attenuator set at a nominal "zero," a 1000-cps signal at standard recording level was played back in one stereo channel from RCA Victor Standard Test Tape No. 12-5-61T. The voltage produced by this signal at the attenuator output, when doubled to allow for the fact that the test tone was reproduced in only one channel whereas the musical program was the addition of signals from two channels, was taken as the standard recording level corresponding to 3.9 cm/sec rms in a single 45/45 stereo channel. This comparison of zero levels on tape and disk has been made arbitrarily for the purposes of the present tests. In commercial-recording practice, there is no standardized relation between the two levels. Further comments on this point will be made in subsequent sections. By means of Eq. (7) the velocity level of a sinusoidal signal which would just reach the curvature-overload limit for a 0.7-mil stylus tip at a 5-in. groove diameter was calculated. For example, curvature overload for $f = 10$ kc, $r = 0.0007$ in. and $D = 5$ in. occurs for $v_{max} = 4.35$ cm/sec, which is 2.1 db below the standard-recording level based on the RIAA equalization characteristic. A signal at this overload level was fed to the measur-

ing system from a signal generator, and the counter threshold was adjusted until the counter just began to register.

After having calibrated the system in this way, musical selections from the stereotapes were played and the number of counts of curvature overload peaks was noted for each selection. Counts of occurrences of overloading were made at level intervals of 1 db over a 10 db range, using the step attenuator to go from +4 db to -6 db relative to the standard level.

Velocity Correction in Measurement of Curvature Overloading

It was pointed out in Section 1c that the maximum groove curvature (i.e., the reciprocal of the radius of curvature) is proportional to the stylus acceleration only for sinusoidal modulation, in which case the maximum curvature is

$$c = 1/V_g^2 \cdot d^2y/dt^2. \quad (9)$$

For a complex waveform the exact expression for the curvature is

$$c = 1/V_g^2 \cdot d^2y/dt^2 [1 + (1/V_g^2 \cdot dy/dt)^2]^{-3/2} \quad (10)$$

and the instantaneous values of both acceleration and velocity must be taken into consideration. In most of the preceding discussion of curvature overloading in this paper, as well as in discussions of this topic elsewhere in the technical literature, sinusoidal signals have been assumed, and curvatures were calculated on that basis. The first point which we wish to investigate here is the error arising when the velocity term in Eq. (10) is taken to be zero when calculating or measuring groove curvatures in the case of musical selections.

The system for counting the occurrences of curvature overloading will be somewhat more complex when the velocity correction is made. A simple differentiating network can provide a voltage proportional to acceleration as indicated in Fig. 5. In order to make the velocity correction, the voltage coming from the differentiator must be modulated in accordance with the function $[1 + (1/V_g \cdot$

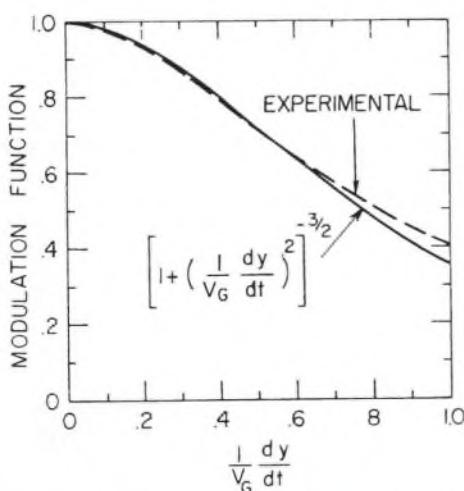


FIG. 6. Modulation function as desired and as used for velocity correction.

$dy/dt)^2]^{-3/2}$ to give a voltage which is proportional to the instantaneous curvature in a complex wave. A vacuum-tube circuit which performs this function with an acceptable degree of precision was devised. In a velocity-corrected counting system, this circuit replaced the differentiator and full-wave rectifier blocks in Fig. 5, since the circuit performed both of these operations as well as providing the required modulation function. By proper adjustment of the operating conditions of the circuit it was possible to achieve a gain-vs-bias characteristic for the modulator which closely approximated the function required for velocity correction over a range great enough to include all values of stylus velocity of interest. The gain-vs-bias characteristic used in the experimental work as well as the desired exact characteristic is plotted in Fig. 6. The over-all gain of the circuit was adjusted so that the output voltage of the circuit was proportional to the curvature in a groove having a 5-in. diameter on a disk rotating at 33 1/3 rpm.

Counts of Curvature-Overload Occurrence

The results of counting the occurrences of curvature overloading for seven musical selections are plotted in Figs. 7

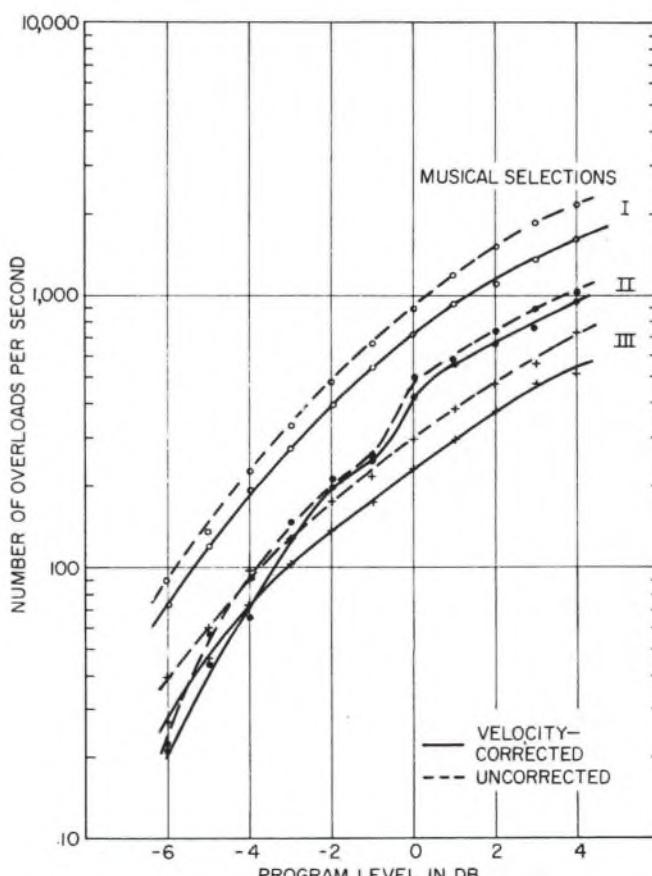


FIG. 7. Average number of occurrences per second of curvature overloading for 33 1/3 rpm, a 0.7 in. stylus tip radius, and a 5-in. groove diameter.

and 8. The total number of occurrences for each selection was divided by the duration of the passage in seconds, and

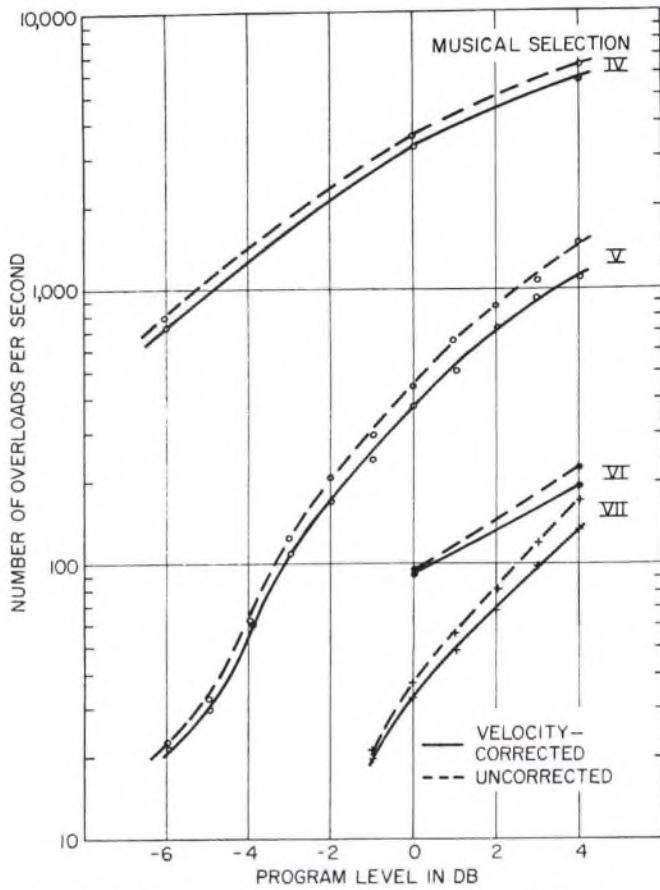


FIG. 8. Average number of occurrences per second of curvature overloading for $33\frac{1}{3}$ rpm, a 0.7-in. stylus tip radius, and a 5-in. groove diameter.

the average number of overloads per second was plotted as a function of the program level in db relative to zero program level as defined in Section 2a. Both uncorrected and velocity-corrected data are shown. We observe that for some types of music a very large number of overloads occur even for zero program level for the conditions specified, i.e., $33\frac{1}{3}$ rpm, 5-in. groove diameter, and a 0.0007-in. pickup stylus tip radius. For most of the musical passages, a decrease of 6 db in level reduces the number of overloads by roughly an order of magnitude.

The velocity correction of the curvature function is seen to reduce the number of overloads by a significant amount. However, the shapes of the overloads/sec-vs-level curves with and without velocity correction are very similar. For most practical purposes the corrected curve may be obtained by shifting the uncorrected curve 1 db in the direction of higher levels. This is a fortunate circumstance since it means that for most purposes curvature overloading may be measured simply in terms of stylus acceleration with a 1 db correction in the program level.

3. MEASUREMENT OF OCCURRENCES OF SLOPE OVERLOAD

The instrumentation used for counting slope overloads

was the same as that shown in Fig. 5, except that the differentiator was replaced by a network having a flat response-frequency characteristic. Calibration of the counter threshold was accomplished by using the 1 kc standard-level signal from RCA Victor Standard Test Tape No. 12-5-61T, as before, and by calculating the slope-overload level using Eq. (4). The conditions specified for these measurements were $D = 5$ in. and $\Omega = 33\frac{1}{3}$ rpm. The average number of slope-overloads per second as a function of program level is plotted for several musical passages in Fig. 9.

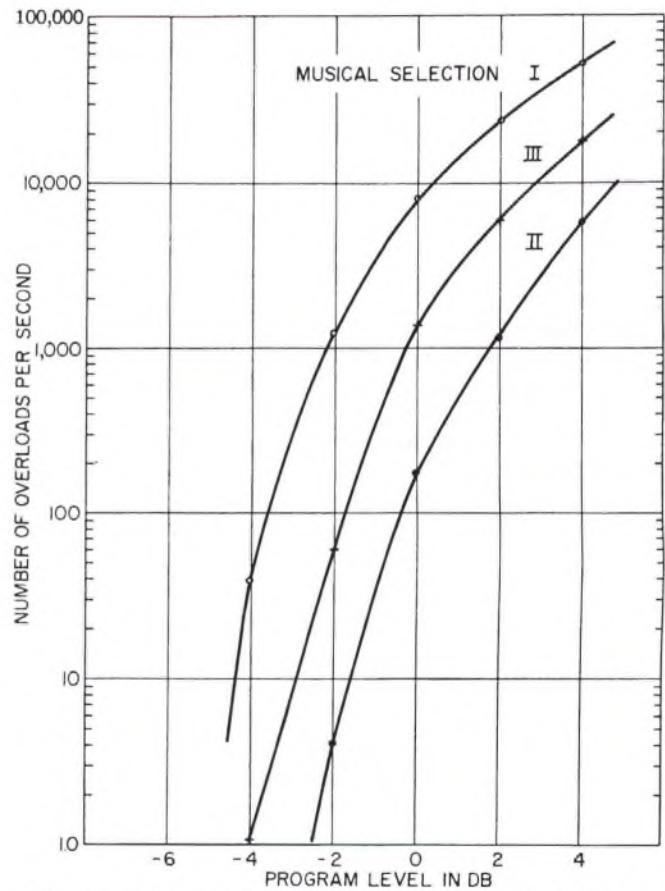


FIG. 9. Average number of occurrences per second of slope overloading for $33\frac{1}{3}$ rpm and a 5-in. groove diameter.

4. MEASUREMENT OF DISPLACEMENT OVERLOADING

The same system shown in Fig. 5 for curvature-overload measurement was used for counting the occurrences of displacement overloading, except that in this case the differentiator was replaced by an integrator. Equation (2) and the Standard Test Tape were used to calibrate the counter threshold to correspond to various values of groove pitch. A typical set of measurements is shown by the solid curve in Fig. 10 for a musical selection having high-level, low-frequency content. The ordinate gives the total number of overcuts which would occur during the entire 87-second passage for the number of grooves per inch indicated on the abscissa.

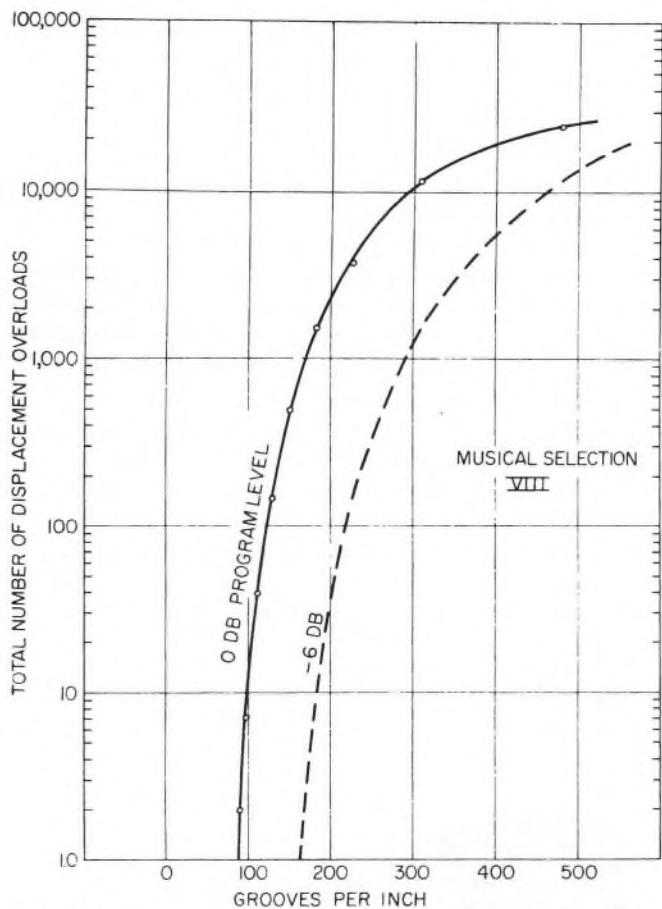


FIG. 10. Measured occurrences of displacement overloading as a function of the number of grooves per inch and for a minimum groove width of 1 mil.

5. CONCLUSIONS

The experimental counts of the occurrence of curvature, slope, and displacement overloading for several musical passages have been plotted in Figs. 7 through 10. The data show the number of occurrences as a function of program recording level relative to an arbitrary zero level. This zero level was established by assuming that the standard recording level on a tape corresponded to a recording stylus peak velocity of 5.5 cm/sec in a stereo channel. As was mentioned earlier, this correspondence does not exist generally in practice. It appears, rather, that every disk recording is a standard unto itself, with the adjustment of the recorded program level being based on subjective judgments not always consistent with objective measurements of overloading. It seems clear, however, that most disk recordings are made at levels noticeably lower than the level arbitrarily taken as zero in this work. The first indication that this is the case is found in the measurements of displacement overloading as plotted in Fig. 10.

Displacement overloading, or overcutting, does not result merely in audible distortion of the reproduced signal but, rather, in groove skipping of the pickup. Hence, even one occurrence of overcutting means rejection of the master disk. The data of Fig. 10 show that in order to avoid over-

cutting in the recording of Selection VIII the groove pitch must be such that fewer than 90 grooves per inch could be accommodated. Such groove displacements are too large to be acceptable in many commercial phonograph systems, and it would be necessary to reduce the program level of this selection by about 6 db. The overload counts which would be obtained if the program were attenuated by 6 db are shown by the dashed curve of Fig. 10.

In order to provide additional information for estimating the relation between the zero level used in these tests and as actually used in disk-recording practice, Selection I was recorded at various levels and at several diameters on a lacquer disk. Playback levels for this disk were compared with those found on commercial pressings. Also, by listening to the program reproduced through a high-quality phonograph system, it was possible to select the recorded level at which the distortion at the inner diameter was comparable to that found on commercial pressings. By means of these tests it was concluded that commercial recordings are made at levels estimated to be from 3 to 8 db below the zero level used in the work reported here. This considerable spread in levels is required to accommodate various types of music and also various judgments in the recording industry as to what constitutes tolerable distortion on a record. In any case, interpretations of the data presented in this paper must take into account the fact that the program level used in recording disk masters will be several db below the zero level indicated in Figs. 7 through 10.

The results for displacement overloading require little additional comment since this overload limit bears a simple, linear relation to program level and since overcutting is easily observed and corrected in practice.

Slope- and curvature-overload limits may be exceeded simultaneously in a record groove. Hence, it is sometimes difficult to separate the two in actual disk-recording tests. The instrumentation used in the present work provides a means for predicting each kind of overload independently, and the measurements of the number of slope overloads as a function of program level offer data not heretofore available. Plots of these measurements result in curves having very similar shapes for the three musical selections which were measured. The number of overloads is observed to decrease rapidly as the program level is reduced. It appears that, for the program levels ordinarily used in cutting master disks, slope overloading is a rather rare event.

Curvature overloads, on the other hand, occur with great frequency for some types of music, and may average as hundreds of occurrences per second for the innermost groove diameters on a record disk. Curvature overloads generally occur in bursts which can be correlated with the high-frequency-energy content of certain musical instruments. Muted trumpets, for example, are an unusually potent source of curvature overloads. These tests have also demonstrated that stylus acceleration can be taken as a measure of the groove curvature, but counts of the number of overloads measured on this basis will be too high and will correspond to the true number occurring for a program level

approximately 1 db higher. This discrepancy results from neglecting the velocity term in the analytical expression for groove curvature.

It is believed that the data presented in this paper give an accurate description of the dependence of each type of overloading on program level. However, these data, alone, do not permit more than tentative conclusions to be made with regard to acceptable recording levels. This is true for two reasons. The first is the arbitrary nature of the zero level of the program material as discussed above. The second cause of uncertainty is our incomplete knowledge of the subjective seriousness of the various types of overloading. Displacement overloading, or overcutting, we know to be almost totally intolerable since it leads to groove skipping or adjacent-groove crosstalk in playback. Consideration of the physical nature of slope overloading suggests that it should result in intermodulation distortion to a degree which is probably quite unacceptable. However, the present writers are not aware that any evaluation of the audible effects of slope overloading in the absence of other complicating distortions has been made. The fact that curvature over-

loading occurs frequently in record releases which are considered acceptable suggests that this form of overload is relatively tolerable. Again, no firm data are available to permit judgments as to the subjective seriousness of curvature overloading. We might hope that experiments will be undertaken to permit a subjective evaluation of slope and curvature overloading as well as other forms of distortion not considered in this paper.

One final point may be mentioned. All of the measurements reported here rest on the basis of the RIAA recording-equalization characteristic. In practice it is customary to use "diameter equalization" which boosts the high frequencies above the levels of the RIAA characteristic as the inside of the record disk is approached during the cutting of a master disk. This is done in order to partially compensate for the tracing loss of short-wavelength, high-frequency signals in playback. This boosting may amount to from 2 to 4 db at 10 kc and its effect is to increase the number of occurrences of slope and curvature overloading. This additional factor must be taken into account in arriving at any final judgment regarding acceptable recording levels.

THE AUTHORS



J. G. WOODWARD



E. C. FOX

J. G. Woodward received his B.S. in 1936 from North Central College, his M.S. in physics in 1938 from Michigan State University, and his Ph.D. in physics in 1942 from The Ohio State University. In March, 1942, he joined the engineering activity of the RCA Victor Division in Camden, New Jersey and later that year moved to the newly formed RCA Laboratories in Princeton, New Jersey, where he has since been engaged in research in audio and electro-acoustics.

Dr. Woodward is a member of the Audio Engineering Society, a Senior Member of the IRE, a member of the IRE Professional Group on Audio, a member of Sigma Xi, and a Fellow of the Acoustical Society of America and of the American Association for the Advancement of Science.

Edward C. Fox received his B.S. and M.S. degrees from Massachusetts Institute of Technology in 1957, as a graduate of the cooperative program. During this program he was employed part-time at the Philco Corporation, where he worked on guided missiles, transistor switching circuits, and an optical phonograph pickup.

After graduation he joined the RCA Laboratories in Princeton, where he has been engaged in research on tape and disk recording.

Program-Level Overloading and Equalization in Phonograph Recording*

E. C. FOX AND J. G. WOODWARD

RCA Laboratories, Princeton, New Jersey

Various modifications were made in the RIAA recording-equalization characteristic, and the resulting changes in the number of displacement, slope, and curvature overloads were measured using the instrumentation and techniques described in a companion paper. Tentative data are presented, and problems requiring further investigation are suggested.

1. INTRODUCTION

THREE types of program-level overloading which may occur in phonograph recording were defined and studied in a companion paper.¹ Displacement overloading occurs for the lower audio frequencies and is characterized by excessive lateral or vertical displacements of the recorded grooves. Slope overloading occurs when the slope of the groove modulation exceeds the slopes of the back faces of the recording stylus. Curvature overloading occurs for the high-frequency components of the recorded program and is characterized by groove modulations having radii of curvature smaller than the tip radius of the playback stylus. By using the instrumentation and techniques described in the companion paper in conjunction with programs reproduced from magnetic-tape recordings, it is possible to count the number of times each type of overload would occur in a disk recording of the program made under specified conditions. We thus have a means for making measurements of the way in which each type of overload depends on the form of recording equalization which is used.

The RIAA equalization characteristic now generally accepted as standard by the recording industry is a compromise which has evolved from numerous antecedent characteristics. The most significant feature of the RIAA characteristic as well as of most of its predecessors is the increasing gain with increasing frequency. This boosting of high frequencies in recording, along with its required inverse characteristic in reproduction, yields a higher signal-to-noise ratio than would be possible with a flat recording character-

istic while maintaining tolerable groove displacements at low frequencies. The boosting of the high frequencies was originally based on the supposition, largely borne out by experimental measurements, that musical and vocal program material carried less energy in the high-frequency portion of the spectrum than in the low-frequency and mid-band portions.

At the present time there is a trend in some segments of the recording industry to record music at ever higher levels in order to make the reproduced program sound as "loud" as possible. At the same time, much of the modern music carries more energy in the high-frequency portions than was true a decade or more ago. The result is frequent occurrences of curvature overloading. The question is raised from time to time as to the advisability of modifying the RIAA characteristic in some way to better accommodate present-day needs and practices. Experiments of the type described in this paper offer a means for an objective appraisal of the effect of changes in the recording characteristic on the number of overloads which occur. Some preliminary tests of this sort were conducted. These will be described in the following sections.

2. EQUALIZATION AND CURVATURE OVERLOADING

Using the curvature-overload-measuring system described previously, the total number of occurrences of overloads in a musical selection** were counted for various equalization characteristics. The counting threshold was adjusted to correspond to a groove diameter of 5 in. on a 33 $\frac{1}{3}$ rpm record, and a pickup-stylus tip radius of 0.0007 in. Two different groups of modifications were made in the RIAA recording equalization for these tests. In the first group various high-

* Presented October 16, 1962 at the Fourteenth Annual Fall Convention of the Audio Engineering Society, New York.

¹ J. G. Woodward and E. C. Fox, "A Study of Program-Level Overloading in Phonograph Recording," *J. Audio Eng. Soc.* 11 (Jan., 1963).

** Ref. 1, Selection I.

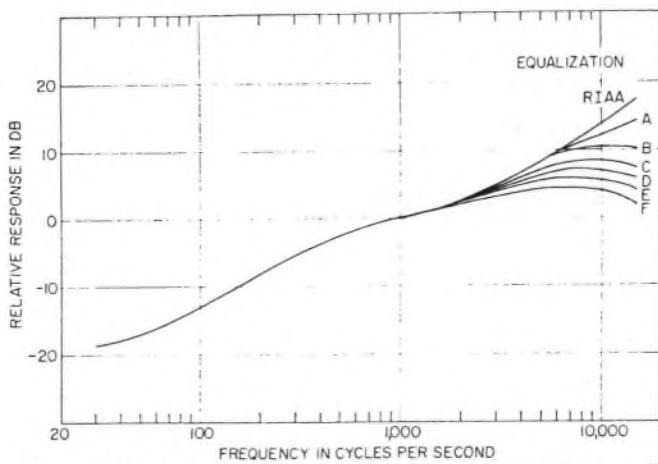


FIG. 1. First group of modifications of the RIAA recording-equalization characteristic.

frequency roll-offs were added to the RIAA characteristics. These characteristics are shown in Fig. 1 as curves *A* through *F*. In the second group of modifications the upper-frequency turnover was shifted to various frequencies, i.e., the time constant which is normally 75 μ sec in the RIAA characteristic was given various shorter values. This group of characteristics is shown in Fig. 2 by curves *G* through *K*. The counts of occurrences of curvature overloads for the musical

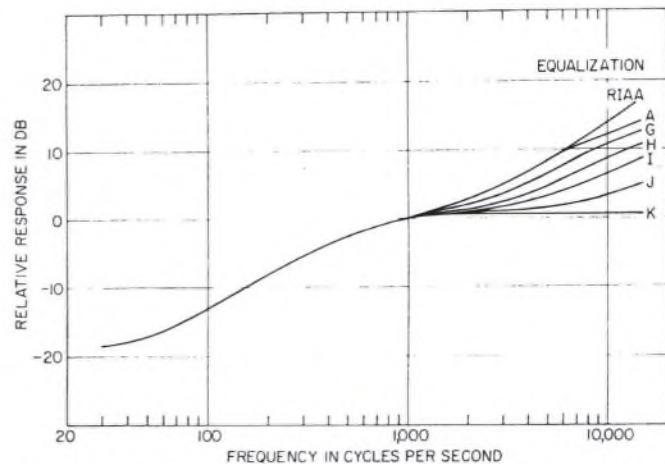


FIG. 2. Second group of modifications of the RIAA recording-equalization characteristic.

selection when the various recording-equalization characteristics are used are plotted in Figs. 3 and 4 as a function of the program level. Due to stray capacitance in the circuits, the over-all response of the system when an RIAA equalizer was used was somewhat low at the high-frequency end of the band. Equalization *A* is the actual measured response when an RIAA network was inserted.

The data of Figs. 3 and 4 show the expected result,

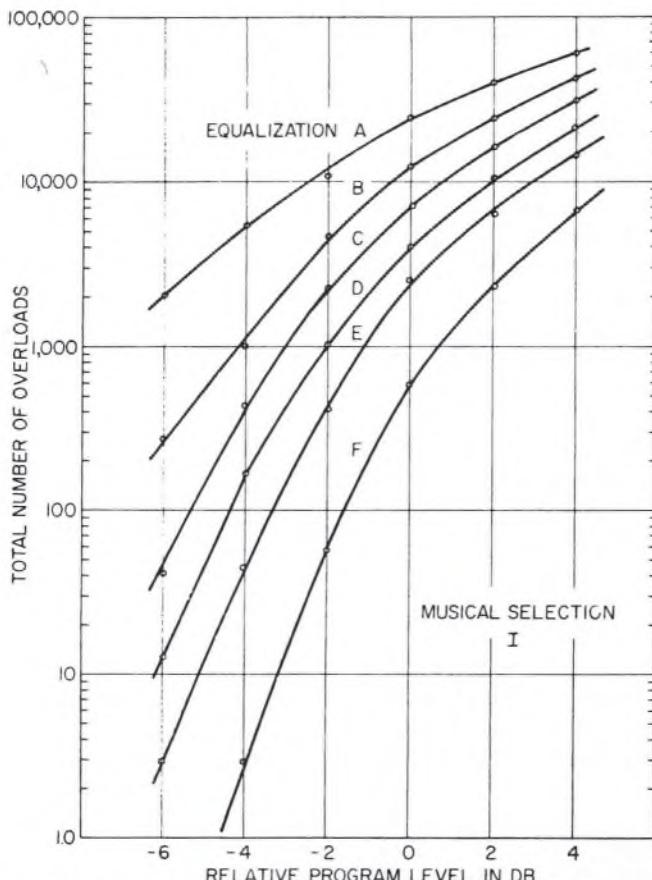


FIG. 3. Measured occurrences of curvature overloading for various recording-equalization modifications.

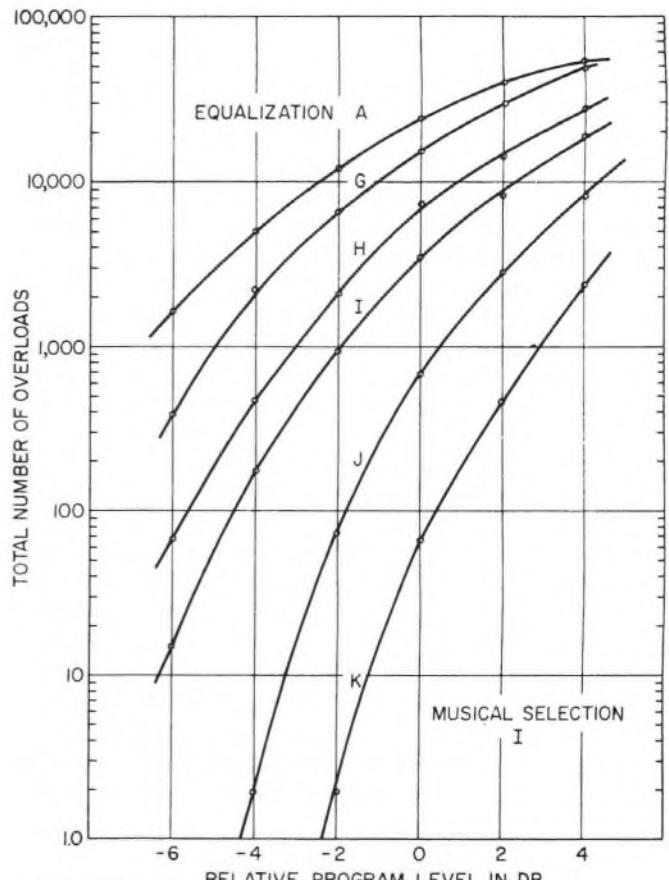


FIG. 4. Measured occurrences of curvature overloading for various recording-equalization modifications.

namely, that reducing the high-frequency boost in the recording equalization produces fewer overloads. We must now attempt to assess the other consequences of the modifications made in the equalization. Principal among these is the change in the signal-to-noise ratio in the reproduction of a recording. In considering the signal-to-noise ratio we will assume a velocity-sensitive pickup having a flat response from 30 to 15,000 cps. We will further assume that the power spectrum of the reproduced surface noise is flat over this bandwidth. Data presented by Howling² as well as measurements made in the present writers' laboratory indicate that the assumption of the noise spectrum flat on a velocity basis is reasonable for present purposes. When a change is made in the recording equalization, a corresponding change must be made in the playback equalization in order to maintain a flat over-all system response. The playback equalization must always be the inverse of the recording equalization. In Figs. 5 and 6 are plotted the playback-

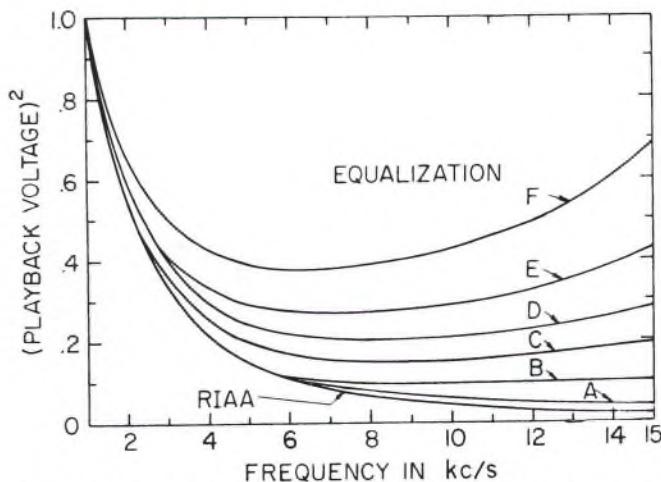
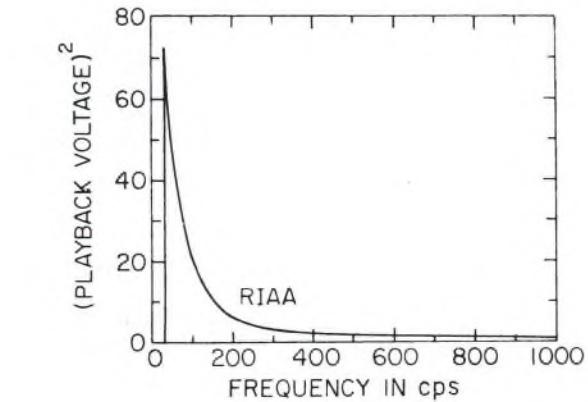


FIG. 5. Playback equalization characteristics plotted for the purpose of calculating relative noise powers.

equalization characteristics corresponding to the recording-equalization characteristics in Figs. 1 and 2. The playback characteristics are plotted against a linear frequency scale

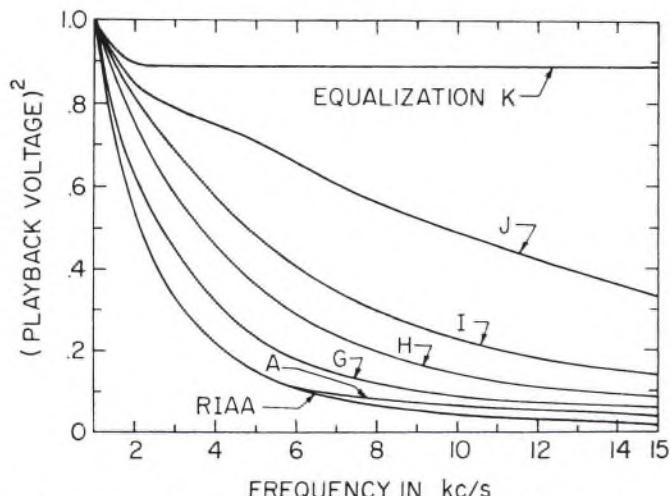


FIG. 6. Playback equalization characteristics plotted for the purpose of calculating relative noise powers.

as abscissa, while the ordinate is on a scale which is proportional to power rather than voltage. When plotted in this way, the area under a playback-equalization curve is proportional to the wide-band noise power for the ideal pickup and noise spectrum assumed. The area under each of the playback-equalization curves in Figs. 5 and 6 was measured. These areas are shown in Tables I and II, with unit area

TABLE I. Area under the playback-equalization curves of Fig. 5.

Equalization	RIAA	A	B	C	D	E	F
Area under curve	1.00	1.08	1.18	1.46	1.75	2.07	2.78
Relative noise power in db	0	0.3	0.7	1.3	2.4	3.2	4.4

TABLE II. Area under the playback-equalization curves of Fig. 6.

Equalization	RIAA	A	G	H	I	J	K
Area under curve	1.00	1.08	1.34	1.75	2.21	3.30	4.90
Relative noise power in db	0	0.3	1.3	2.4	3.5	5.2	6.9

being taken as the area under the RIAA curve. These values, when converted to decibels, also give the relative wide-band noise power in playback for each equalization.

The data presented in Figs. 3 and 4 and Tables I and II allow us to make and assess various compromises between signal-to-noise ratio and curvature overloads. As an example let us compare system performance with equalization characteristics A and F. If we are to keep the same signal-to-noise ratio in each case, we must record at a 4.1 db higher level when using F than when using A. When measuring at zero level with A, approximately 24,000 curvature overloads occurred, as shown in Fig. 2. For a level of +4.1 db and characteristic F, the number of overload counts was 6,800. In a similar way, compare the results when a level of -6 db is used with A and of -1.9 db with F. The overload counts are now 21,000 and 68, respectively. These examples indicate that modifications of the RIAA recording equalization characteristic of the type used in the first group, i.e., A

² D. H. Howling, *J. Acoust. Soc. Am.* 31, 1463 (November, 1959).

through F , can result in a significantly lower number of occurrences of curvature overloading while maintaining the same over-all response-frequency characteristic and signal-to-noise ratio.

Now let us take examples from the second group of modifications and compare, first, the results of using A at zero level and J at a level of 4.9 db, which would give the same signal-to-noise ratio in both cases. The counts of curvature overload shown in Fig. 4 are 24,000 and 13,500, respectively. When using A at -6 db and J at -1.1 db, the counts are 1600 and 230, respectively. Again, the modification results in fewer occurrences of curvature overload. However, the modifications of the second group must be much more drastic than those of the first group in order to produce the same reduction in the number of overloads.

It has just been shown that the number of curvature overloads may be reduced significantly by modifying the high-frequency portion of the recording-equalization characteristic. When this is done, the program level must be increased if the signal-to-noise ratio is to be kept the same. We must now consider whether or not this method of reducing the number of curvature overloads increases the number of slope and displacement overloads significantly.

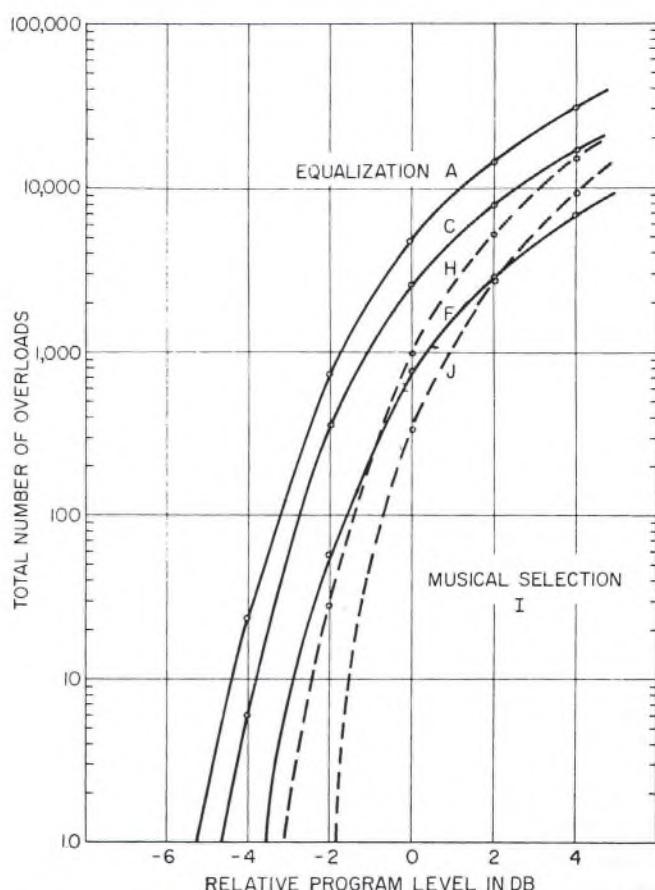


FIG. 7. Measured occurrences of slope overloading for various recording-equalization modifications.

3. EQUALIZATION AND SLOPE OVERLOADING

Count of the total number of occurrences of slope overloading in musical selection I were made for equalization A , C , F , H and J . The results of these tests are plotted as a function of program level in Fig. 7. Using the same example as in the preceding section, we compare the results of using equalizations A and F . With A at zero level the number of slope overloads was 4800. For F at +4.1 db the number was 7000. For A at -6 db no instances of slope overload were observed. For F at -1.9 db the count was 61. Thus, for both values of recording level, equalization F causes more slope overloads than does A when the signal-to-noise ratio is kept constant.

As the second example, let us compare equalizations A and J . For A at zero level and J at +4.9 db the counts were 4800 and 14,000, respectively. For A at -6 db and J at -1.1 db the counts were zero and 38, respectively. The use of equalization J also results in an increased number of slope overloads for both values of program level. Evidently, for these program levels, it is necessary to trade slope overloads for curvature overloads. However, if one could sacrifice only 2 or 3 db in signal-to-noise ratio by taking the zero level for equalization F or J at -4 db on the level scale used in these tests, both slope and curvature overloads would be almost non-existent in the musical selection used in these tests.

4. EQUALIZATION AND DISPLACEMENT OVERLOADING

None of the equalization modifications discussed above involved the low-frequency portion of the audio band, and hence would not affect the number of occurrences of displacement overloading. Hence, if one of the modifications, such as F , were used in order to reduce curvature overloading, and it were necessary to record at a +4.1 db level the groove displacements would be unacceptably large. However, if recording were done at levels 6 to 8 db lower, as appears to be the case in actual practice, displacement over-

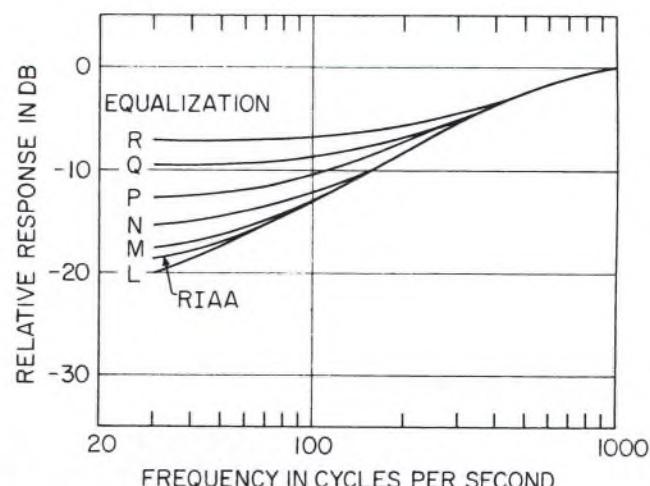


FIG. 8. Low-frequency modifications of the recording-equalization characteristic.

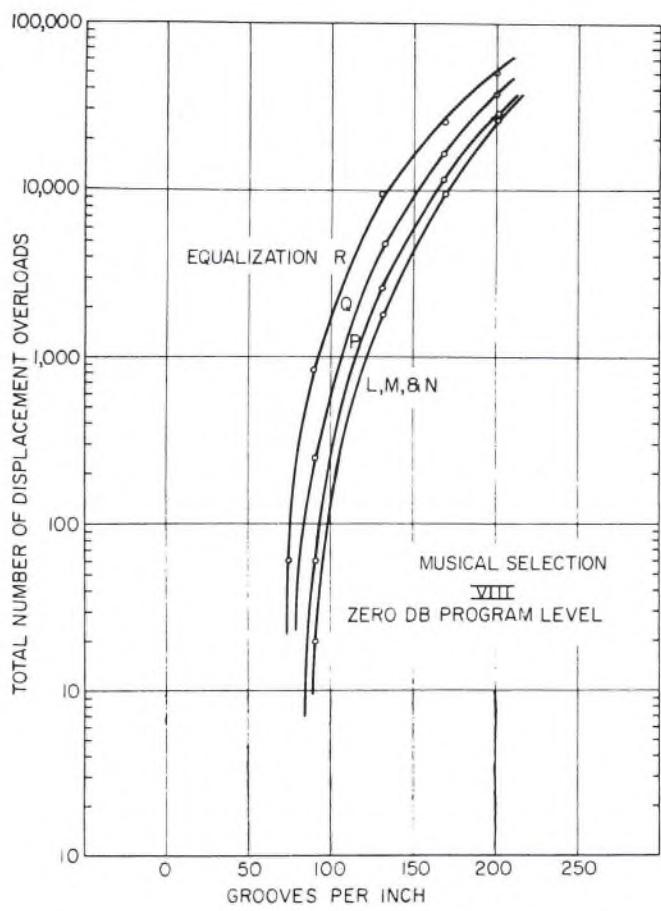


FIG. 9. Measured occurrences of displacement overloading for various equalization modifications.

loading would not be troublesome, and the playing time available on a disk would be reduced by only a few percent.

Several modifications of the low-frequency portion of the recording equalization characteristic were made and displacement overloading was measured for each modification. This group of characteristics is plotted in Fig. 8. Counts of displacement overloads as a function of the number of grooves per inch and for the various equalization characteristics are plotted in Fig. 9 for a musical selection[†] having strong low-frequency components.

Another simpler, and in some ways more useful, measurement of displacement overloading consists of noting the maximum displacement occurring during an entire musical selection. Measurements of this kind were made by means of a peak-reading voltmeter designed to have a charge time of $\frac{1}{3}$ msec and a discharge time of several minutes. When this meter was connected in place of the counter as used in the preceding tests, the meter reading at the conclusion of a musical selection represented the largest displacement peak occurring anywhere in the selection. The meter was calibrated in terms of the 1 kc standard recording level on RCA Victor Standard Test Tape, No. 12-5-61T. Data of

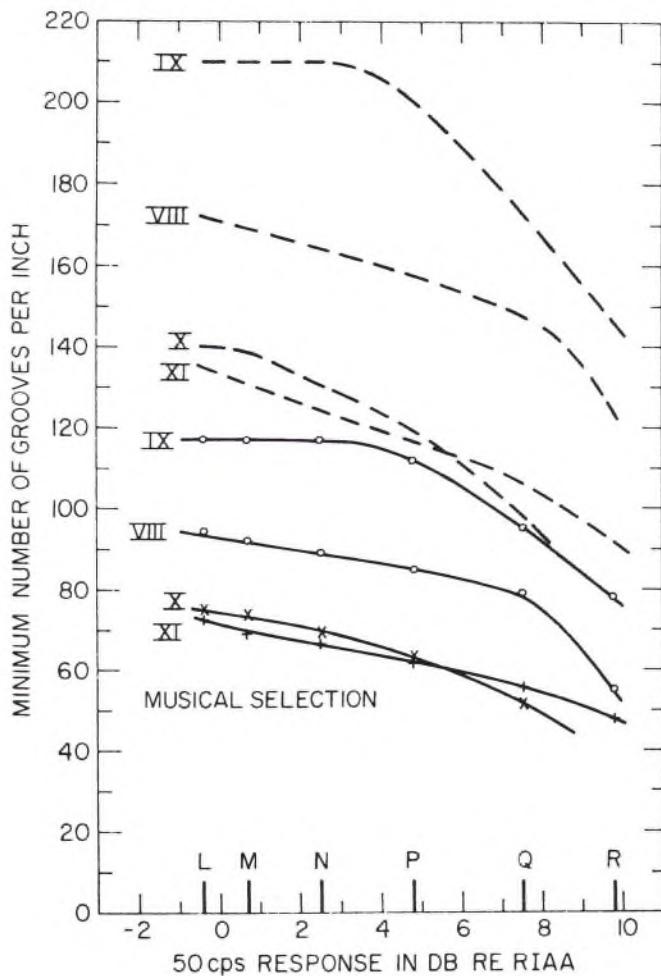


FIG. 10. Minimum number of grooves per inch which can accommodate the largest measured displacement, as a function of recording-equalization modifications, and for several musical selections.

this type were taken for several musical selections[‡] and for each of the low-frequency equalization characteristics shown in Fig. 8. The results are plotted in Fig. 10 where ordinate values are given in terms of the minimum number of grooves per inch for which overcutting will not occur, as determined from the meter readings and Eq. (1) in Ref. 1. The scale of the abscissa is a measure of the difference between the various equalization characteristics, *L* through *R*, being given as the difference in response in db between each characteristic and the RIAA characteristic, this difference being measured at 50 cps. The solid curves in Fig. 10 show the measured data for zero db program level. The dashed curves were calculated from the measurements to present the corresponding information for a program level of -6 db. On the basis of these results one might judge that the low frequencies could be boosted somewhat more than is done in the RIAA recording equalization without paying too high a price in reduced playing time. This would reduce hum and rumble noise in playback. However, it would also accentuate low-frequency tracking problems in the case of pickups having an insufficiently-high compliance.

[†] Ref. 1, Selection VIII.

[‡] Ref. 1, Selections VIII through XI.

5. CONCLUSIONS AND DISCUSSION

Instrumentation employing an electronic counter with a calibrated threshold may be used for measuring displacement, slope and curvature overloading individually. By this means it is possible to study the way in which each type of overload is influenced by modifications of the recording equalization and thus to make objective assessments of various equalization characteristics which may be suggested. The limited number of measurements reported here give some indications of the dependence of overloading on equalization characteristics. However, any conclusions can be only very tentative until many more data for all types of program material have been compiled. One not-unexpected indication is that curvature overloading is most sensitive to the highest-frequency components of the program, i.e., to energy carried by components above about 3 kc/s. Slope overloading is also sensitive to energy in this frequency range but is, in addition, affected by energy in the band at least down to 1 kc/s. Modifications which reduce the gain in the high-frequency region of the recording equalization act to reduce the number of curvature overloads significantly even when the program level is increased in order to maintain the same signal-to-noise ratio in playback. However, this modification and increased level result in a greater number of slope overloads. Because of the rapid rate of change of the number of slope overloads as a function of program level, it appears feasible to reduce the occurrence of both slope and curvature overloads to negligible values. The price for this improvement in quality would be, as expected, a reduction in the available playing time on a disk in order to avoid displacement overloading, and also a slightly lower signal-to-noise ratio.

Modifications of the low-frequency end of the recording-equalization characteristic affect the occurrence of displacement overloading. Measurement of displacement overloads

for several low-frequency modifications suggest that somewhat greater boosting of frequencies below about 100 cps can be achieved without decreasing the number of grooves per inch. However, it is unlikely that the small advantage to be gained in reduced low-frequency noise in playback justify this change in the recording equalization standard, particularly since pickup tracking problems would be accentuated.

It should be recalled that the curvature- and slope-overload counts reported here are based on instrumentation calibrations corresponding to a 5-in. groove diameter, i.e., the innermost grooves on a $33\frac{1}{3}$ rpm, 12-in. record. At larger diameters, overloading is less frequent. While the techniques described and illustrated in this work permit useful objective measurements of overloading, they are not a completely sufficient basis for standardizing levels and equalization characteristics in disk recording. The overload limits are arbitrary, being defined in terms of simple mathematical expressions and geometrical configurations, but with the knowledge that significant amounts of distortion occur even before these limits are exceeded. Some subjective judgements are required to establish a relationship between the overload limits and a tolerable value of distortion.

The tests reported here suggest that it may be necessary to trade curvature overloads for slope overloads when a recording-equalization characteristic is modified. It is desirable, therefore, to have subjective judgements as to the relative seriousness of these two forms of distortion. It is hoped that experiments can be devised which will permit individual adjustment of slope overloading and playback tracing distortion (which progresses into curvature overloading), and which can yield listener judgements of tolerable levels of each form of distortion. Work along these lines can be expected to provide important, additional information and incentive for improvement and standardization in phonograph recording systems.

Tracing Distortion—Its Cause and Correction in Stereodisk Recording Systems

E. C. FOX AND J. G. WOODWARD*

RCA Laboratories, Princeton, New Jersey

The phase-modulation effects which are responsible for tracing distortion in stereodisk systems are discussed, and analytical expressions are provided for calculating the magnitudes of the theoretical distortion products. Experimental results for the reduction of tracing distortion by re-recording techniques are presented and a method for electrically simulating tracing distortion is described. Simulation techniques have been embodied in a device known as the Dynamic Recording Correlator; results of using it in a recording system to reduce tracing distortion in playback are presented.

STEREODISK masters are cut with a stylus having sharp edges, which makes it possible to record very-short-wavelength modulations on the groove walls. The recorded signals are played back by means of pickups fitted with styli having spherical tips of finite dimensions. The disparity between the shapes of the recording and playback styli results in distortion in the reproduced signal. This distortion is termed "tracing distortion" because it has its basis in the geometrical relationships between the waveform of the recorded modulation and the spherical stylus tip used to trace it in playback. The geometrical basis of tracing distortion has been known for many years. DiToro¹ provided an approximate analysis of tracing distortion for vertical recording. This was followed by progressively more exact analyses for both vertical and lateral recording by Pierce and Hunt,² Lewis and Hunt,³ and Corrington.⁴

Among other things, these analyses showed, and measure-

ments verified, that tracing distortion was much lower for lateral modulation than for vertical modulation. As long as the great bulk of commercial records were laterally modulated there was little need for further study of the problem of tracing distortion. With the advent and large-scale commercial acceptance of 45/45 stereodisk records, tracing distortion again became a significant limiting factor in the quality of the reproduced sound, since each channel of the 45/45 system is analytically equivalent to a vertically modulated groove. The harmonic and intermodulation distortions in 45/45 stereodisk systems have been calculated by Corrington and Murakami.⁵

The present study is both analytical and experimental. In the following section the mechanism involved in producing tracing distortion will be described, and expressions will be given for calculating both harmonic and intermodulation distortion. Subsequent sections will deal with techniques for reducing tracing distortion. These techniques are based on the complementary nature of tracing distortion and include actual re-recording experiments as well as electronic simulation of tracing distortion.

THE MECHANISM OF TRACING DISTORTION

Our present concern is with the tracing distortion gener-

* The biography of the two authors appears in the January 1963 issue of the *Journal* on p. 23.

¹ M. J. DiToro, "Distortion in the Reproduction of Hill and Dale Recording," *J. Soc. Mot. Pict. Engrs.* 29, 493 (1937).

² J. A. Pierce and F. V. Hunt, "Distortion in Sound Reproduction from Phonograph Records," *J. Soc. Mot. Pict. Engrs.* 31, 157 (1938).

³ W. D. Lewis and F. V. Hunt, "Theory of Tracing Distortion in Sound Reproduction from Phonograph Records," *J. Acoust. Soc. Am.* 12, 348 (1941).

⁴ M. S. Corrington, "Tracing Distortion in Phonograph Records," *RCA Review* 10, 241 (June, 1949).

⁵ M. S. Corrington and T. Murakami, "Tracing Distortion in Stereophonic Disc Recording," *RCA Review* 19, 216 (Jan., 1958).

ated in playback of a single channel of a 45/45 stereodisk. Since it is not possible to derive analytical expressions in closed form for the distortion components, calculations of the magnitudes of these components are based on series approximations. The exactness of the resulting calculations is dependent on the number of terms retained in the series. Corrington's⁴ analysis has been extended to yield the following expressions for the second-harmonic and first-order-intermodulation products for a 2-component signal recorded on one channel of a stereodisk.

For a recorded modulation of $a_1 \sin 2\pi f_1 t + a_2 \sin 2\pi f_2 t$, the amplitudes of components in playback are as shown in Table I.

TABLE I. Amplitudes of components in playback for a recorded modulation of $a_1 \sin 2\pi f_1 t + a_2 \sin 2\pi f_2 t$.

Frequency	Amplitude
f_1	$a_1 \left\{ 1 - \frac{k_1^2 r^2}{4} \left(\frac{k_1^2 a_1^2}{2} + k_2^2 a_2^2 \right) \right\}$
f_2	$a_2 \left\{ 1 - \frac{k_2^2 r^2}{4} \left(\frac{k_1^2 a_1^2}{2} + \frac{k_2^2 a_2^2}{2} \right) \right\}$
$2f_1$	$\frac{k_1^2 a_1^2 r}{4} \left\{ 1 - \frac{k_1^2 a_1^2}{4} - \frac{3k_2^2 a_2^2}{4} - k_1^2 r^2 \left(\frac{k_1^2 a_1^2}{3} + \frac{k_2^2 a_2^2}{3} \right) \right\}$
$2f_2$	$\frac{k_2^2 a_2^2 r}{4} \left\{ 1 - \frac{3k_1^2 a_1^2}{4} - \frac{k_2^2 a_2^2}{4} - k_2^2 r^2 \left(\frac{k_1^2 a_1^2}{3} + \frac{k_2^2 a_2^2}{3} \right) \right\}$
$f_1 - f_2$	$\frac{k_1 k_2 a_1 a_2 r}{2} \left\{ 1 - \left(\frac{k_1^2 a_1^2 + k_2^2 a_2^2}{8} \right) \left[3 + r^2 (k_1 - k_2)^2 \right] \right\}$
$f_1 + f_2$	$\frac{k_1 k_2 a_1 a_2 r}{2} \left\{ 1 - \left(\frac{k_1^2 a_1^2 + k_2^2 a_2^2}{8} \right) \left[3 + r^2 (k_1 + k_2)^2 \right] \right\}$

In these expressions a_1 and a_2 are the displacement amplitudes of the two recorded signal components of f_1 and f_2 , respectively; $k_1 = 2\pi f_1 / V_G$ and $k_2 = 2\pi f_2 / V_G$, where V_G is the linear groove velocity; and r is the tip radius of the playback stylus. Because of the inclusion of a greater number of terms, these expressions are somewhat more accurate than the single-term approximation given by Corrington and Murakami⁵ for the intermodulation components. The amplitudes given by these expressions apply to the case of a displacement-responsive pickup. The relative amplitudes for a velocity-responsive pickup may be obtained by multiplying the displacement amplitudes of the components by their respective frequencies.

While series approximations permit the calculation of tracing distortion products, they give little or no insight into the physical mechanism which is responsible for the distortion. The mechanism will, therefore, be considered in more qualitative terms now. Figure 1 depicts a sectional view of a stylus tip and a groove wall in the plane of modulation of one of the 45/45 channels. A sinusoidal modulation is used in this familiar illustration. It is clear that, as the modulated groove moves past the stylus, the point of contact between stylus and groove moves back and forth around the stylus tip. Only when the slope of the modulation is zero

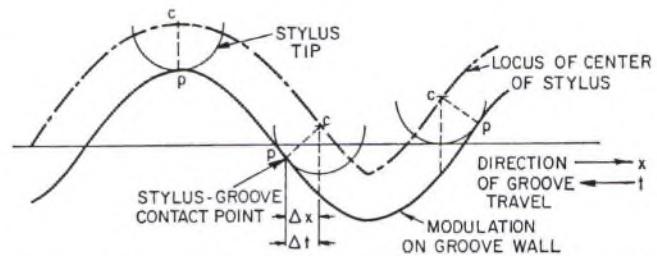


FIG. 1. Representation of a spherical stylus tracing a sinusoidal modulation in the plane of modulation of one of the 45/45 channels.

is the contact point, P , directly below the center of the stylus tip at C . Because of this variation in the location of the contact point, the locus of the center of the stylus as a function of time, or distance along the groove, is a distorted version of the modulation in the record groove. It should be realized, therefore, that the distortion in the pickup out-

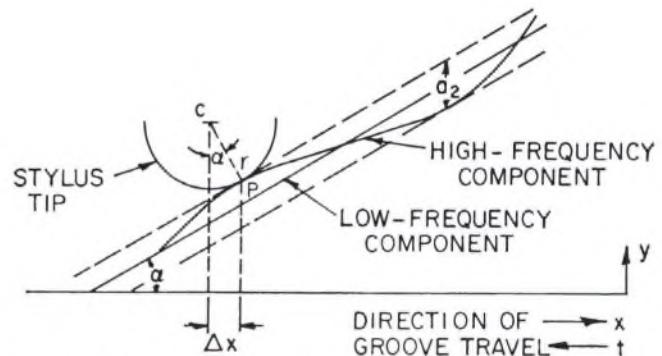


FIG. 2. Representation of a stylus tracing a short section of groove wall modulated with a high-frequency and a low-frequency signal superimposed.

put is not due to a nonlinear transfer characteristic in the usual sense but, rather, to phase modulation of the signal.

The phase-modulation basis of tracing distortion is especially evident in the case of a 2-component recorded signal for which intermodulation products are generated. Such a case is depicted in Fig. 2 for a high-frequency component added linearly to a low-frequency component in recording. The amplitude of the high-frequency component is small compared to the amplitude of the low frequency component, and only a short, positively-sloping, segment of the low-frequency component is shown in the figure. As before, the figure represents a section in the modulation plane of one of the 45/45 channels on the disk. When the low-frequency and the high-frequency components of the groove modulation are

$$y_{LF} = a_1 \sin(2\pi x/\lambda_1)$$

$$y_{HF} = a_2 \sin(2\pi x/\lambda_2),$$

respectively, the instantaneous displacement of the center of the stylus tip is given by

$$y_c = a_1 \sin[2\pi(x + \Delta x)/\lambda_1] +$$

$$a_2 \sin[2\pi(x + \Delta x)/\lambda_2] + (\Delta x / \tan \alpha).$$

To a first approximation, $\tan \alpha$ is the slope of the low-frequency component, and

$$\Delta x = r \sin \alpha = r \sin \tan^{-1}(dy_{LF}/dx) \\ \approx r(dy_{LF}/dx) = (2\pi a_1 r/\lambda_1) \cos(2\pi x/\lambda_1).$$

Retaining only the high-frequency component in playback,

$$y = a_2 \sin \left[\frac{2\pi}{\lambda_2} \left(x + \frac{2\pi a_1 r}{\lambda_1} \cos \frac{2\pi x}{\lambda_1} \right) \right] \\ = a_2 \sin \left[2\pi f_1 t + \frac{4\pi^2 a_1 r f_1 f_2}{V_g^2} \cos 2\pi f_1 t \right].$$

The last equation shows that the intermodulation sidebands due to tracing distortion in a displacement-responsive pickup are caused by the phase modulation of the high frequency by the low frequency. No mechanism is evident to produce amplitude modulation in this case. A velocity-responsive pickup will yield amplitude modulation as well as phase modulation. In actual practice the playback operation contains other nonlinear processes, principally the deformation of the groove wall at the point of stylus-groove contact, which also result in amplitude modulation of the high by the low frequency. The distortion products arising from the various sources combine in complex ways. Shiga,⁶ who has included the groove-wall deformation in an analysis of tracing distortion, and Walton⁷ in an experimental study found that harmonic distortion was reduced as a result of the deformation. These studies did not extend to the case of IM products in the presence of deformation. With currently-available phonograph systems, tracing distortion appears to be more serious than the distortions due to nonlinear effects, and a considerable improvement in the quality of reproduced sound should be possible through the reduction of tracing distortion alone. Methods for achieving such a reduction will now be considered.

CORRECTION OF TRACING DISTORTION

Re-recording Method

It has been known for many years (see References 2 and 3) that tracing distortion could, in principle, be reduced by a re-recording method because the distortion function is complementary. The method involves recording the signal on a disk, using a pickup to play back the recorded signal from the disk, and recording this reproduced signal in reversed phase on a second disk. The signal on the second disk contains the tracing distortion resulting from the first playback, but this distortion is in the correct phase to cancel the distortion occurring during playback of the second disk. The re-recording method was used in former years in the preparation of vertical-cut transcription records. That the technique is effective in stereodisk records also is demonstrated by data shown in Fig. 3. To obtain these data an

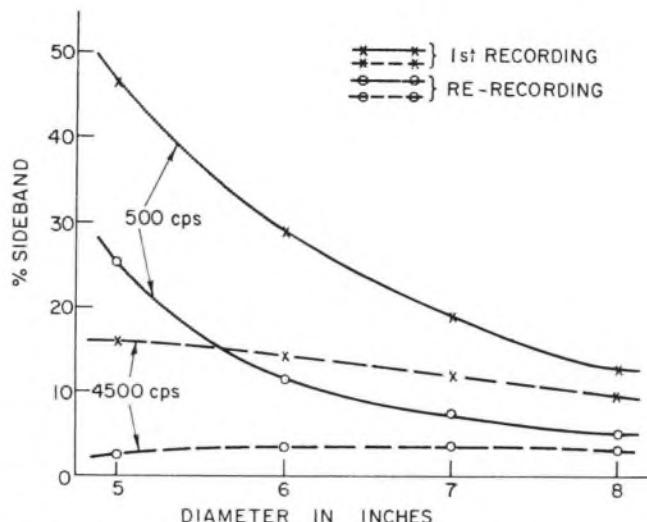


FIG. 3. Intermodulation sidebands measured in playback of a 2000 + 2500 cps signal in a first recording and in a re-recording of the reproduced signal from the first recording.

intermodulation test signal consisting of 2000 + 2500 cps in a 1:1 velocity ratio was recorded in the left-hand channel of a disk at various diameters. This disk was played back and simultaneously re-recorded, in reversed phase, on a second disk. Each recording was then played back and the signal components occurring at 500, 2000, 2500 and 4500 cps were measured by means of a wave analyzer. A velocity-responsive pickup was used. The output of the pickup was passed through an integrator before being fed to the wave analyzer. The amplitude of each sideband was divided by the average of the amplitudes of the 2000 and 2500 cps components to give the % sideband value as plotted. Since distortion due to a vertical tracking-angle error is also complementary, the re-recording technique should be effective in reducing this form of distortion also. However, for the signal used in the tests presented here tracing distortion predominates.

It is clear that a significant reduction of the intermodulation products has been achieved through the re-recording process. It is likely that an even greater reduction could be obtained if the phase and amplitude response of the pickup and playback system were more closely controlled. The difficult practical problem of maintenance of the playback system is one reason why the re-recording technique has not found favor in the commercial recording of stereodisks. Other reasons are the small, but inevitable, increase in noise which results from re-recording, and the very considerable amount of additional time required to make a master lacquer record by the two-stage process. It appears that the complementary nature of tracing distortion can be used to real practical advantage in the reduction of distortion only if an accurate electrical simulation of the signal resulting from the first playback step of the re-recording process can be provided. A device of this sort could be used to process the signal in an otherwise normal stereodisk recording channel so that the signal recorded on the master disk would be equivalent to that resulting from the re-

⁶ Takeo Shiga, "Distortion in Stereo Disk Record (I)," *J. Acoust. Soc. Japan* 18, 118 (1962).

⁷ J. Walton, "Stylus Mass and Distortion," *Wireless World* 69, 171 (April, 1963).

recording process, but without the disadvantages of the latter. The remainder of this report will be devoted to the description of such a simulator and its performance.

Electrical Simulation of Tracing Distortion

The analysis of an earlier section shows that tracing distortion does not arise from a simple nonlinear transfer characteristic but, rather, is due more basically to a phase-modulation process. The process will be discussed further with reference to Fig. 4, which depicts a sectional view of a spherical stylus tip in contact with the modulation on a groove wall of a record. The view as shown is in the plane of the modulation in one of the 45/45 channels on a stereo-disk. A rigid record material is assumed, so the stylus does not deform the material in the region of stylus-groove con-

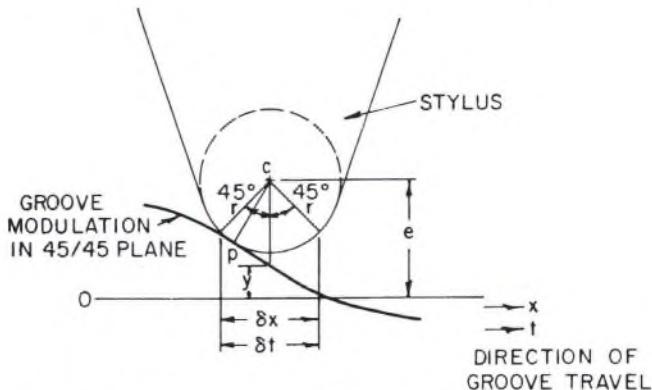


FIG. 4. Representation of a stylus in contact with a modulated groove in the plane of modulation of one of the 45/45 channels.

tact, indicated at *P*. The horizontal axis represents zero displacement, i.e., an unmodulated groove. The stylus moves from right to left along the groove. The instantaneous modulation displacement below the center of the stylus is *y*. For a displacement-responsive pickup the output voltage will be proportional to the vertical displacement (in the plane of the figure) of the center of the stylus, *C*. The radius of the stylus tip is *r*.

If the output voltage is to be an undistorted representation of the groove modulation, the stylus-groove contact point must always be directly below the center of the stylus. In actual fact, the contact point is, in general, either to the right or to the left of the center. Because of limitations in the recording process the slope of the modulation will never (or ought never) exceed $\pm 45^\circ$. Hence, the stylus-groove contact will always fall within the span

$$\delta x = 2r \sin 45^\circ = 1.414r.$$

This span corresponds to a time segment,

$$\delta t = \delta x / V_g = 1.414r / V_g,$$

where V_g is the linear velocity of the stylus along the groove. In the system described here for electrically simulating the output of a pickup tracing the groove modulation, a signal voltage proportional to the modulation displacement is provided. This signal is then displayed and sampled over a

segment of time, δt , as determined above in terms of the stylus size and the stylus-groove velocity. One means of accomplishing this display and sampling process is a tapped delay line. If the total delay time of the line is δt , a segment of the signal of this duration will be present in the line at any given instant. This segment of the signal may be sampled at various points along the line as indicated in Fig. 5a and b. For purposes of illustration, five equally-spaced sampling points are shown.

In this electrical analog of the playback process the shape of the stylus tip may be represented by an applied dc voltage which is adjusted to have different values at different positions along the delay line. Corresponding to the spherical shape of the stylus tip (circular in section in the plane of modulation in the record groove), the "stylus-shape voltage" conforms to a circular voltage-vs-length characteristic along the length of the line. This is illustrated in Fig. 5c. The

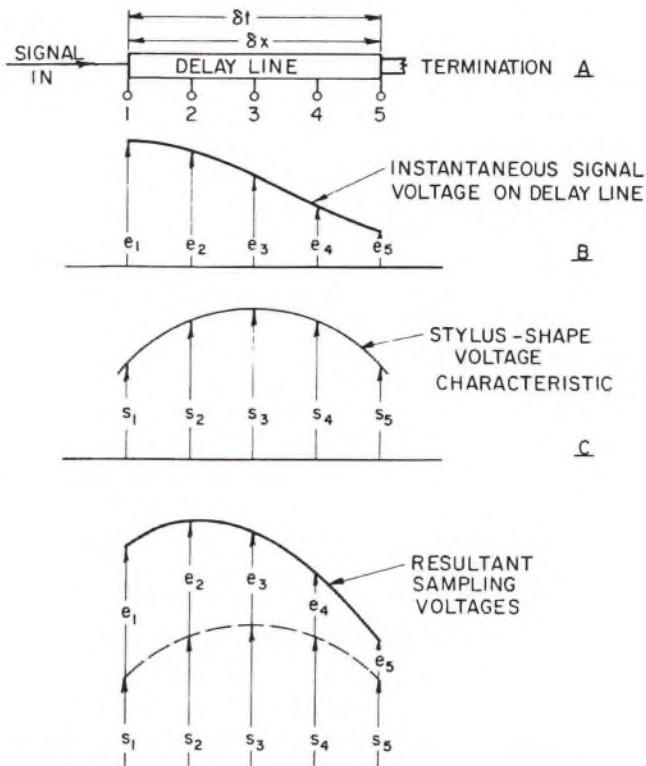


FIG. 5. Diagrammatic representation of the use of a tapped delay line for simulating tracing distortion.

instantaneous signal voltage at each sampling point is added to the stylus-shape voltage to give the resultant sampling voltages as shown in Fig. 5d. Sampling Point 3 corresponds to a stylus-groove contact point directly below the center of the stylus in actual playback. When tracing a modulation waveform as depicted in Fig. 5b, the contact point will be to the left of the center of the stylus and at a position corresponding to Sampling Point 2, within the limits of the present 5-point approximation. The displacement of the center of the stylus and, hence, the instantaneous voltage generated by the pickup will be proportional to $s_2 + e_2$. For other modulation waveforms the stylus-groove contact

will move to other points on the stylus, and there will be a corresponding shift of the maximum value of $S_n + e_n$ to other sampling points in the electrical analog. In every case, however, the maximum value of $S_n + e_n$ occurring at any sampling point in the analog is proportional to the instantaneous value of the voltage at the terminals of a displacement-responsive pickup when reproducing the same modulation waveform. This relationship is only approximate since the analog, as described here, attempts to represent a continuous function by a finite number of sampling points. However, the approximation can be made as exact as one wishes by using a sufficiently large number of points.

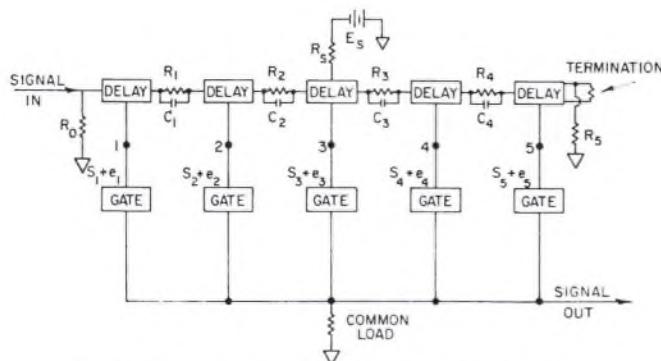


FIG. 6. Block diagram of a tracing-distortion simulator.

One final feature must be provided to complete the electrical analog of the playback process. This feature is a means for sensing and selecting, at every instant, the maximum value of $S_n + e_n$ appearing at the sampling points. This can be accomplished by a series of amplitude-selecting gates. Each sampling point on the line is connected to such a gate. The outputs of all of the gates are connected to a common load. The gates are designed and interconnected in such a way that, in general, at any instant all gates will be closed except the one which is receiving the maximum signal. This gate will pass this signal to the load, as being proportional to the output of a pickup.

With continuously-varying modulation waveforms there will be instants of transition when two gates receive the same value of sampling voltage and, hence, will simultaneously pass equal voltages to their common load. The gate circuits should be so devised that if two gates are passing equal signals simultaneously, each gate which is open loads the other open gate with an impedance equal to its internal impedance. Thus, the resulting voltage across the common load remains the same whether one or two gates are open to the same sampling voltage. This feature removes the possibility of the generation of transient pulses as the maximum value of $S_n + e_n$ moves from one sampling point to another while the signal moves down the delay line with time. The only other occasion on which two gates can receive equal voltages is under conditions of curvature overload, i.e., when the curvature of the signal waveform exceeds the curvature of the stylus-shape-voltage characteristic.

A block diagram of the complete electrical analog of the playback process is shown in Fig. 6. Again, for purposes of

illustration, only five sampling points are used and the delay line is broken into five equal segments. A network including E_s , R_s and R_0 through R_5 indicates the source and adjustment of the dc stylus-shape voltages at each sampling point. The capacitors, C_1 through C_4 , bypass the resistors for audio frequencies, so the delay line is not affected by the presence of the resistors between the segments of the line. The amplitude-selecting gates and their interconnections and common load are also indicated diagrammatically.

THE DYNAMIC RECORDING CORRELATOR

The foregoing principles have been incorporated in a practical electronic device, called a "Dynamic Recording Correlator," for simulating the playback process in a stereodisk system. The Correlator processes the signals in the right-hand and left-hand channels independently with two separate 12-section delay lines and their associated circuitry. A front-view photograph of the device is shown in Fig. 7. The lower, rack-mounting chassis contains the power supply. The upper chassis contains the delay lines, the amplitude-selecting gates, and such amplifier stages and equalization networks

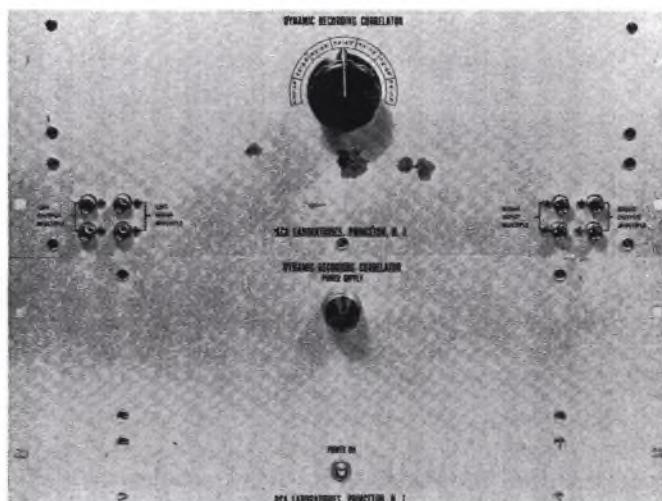


FIG. 7. Front view of the Dynamic Recording Correlator.

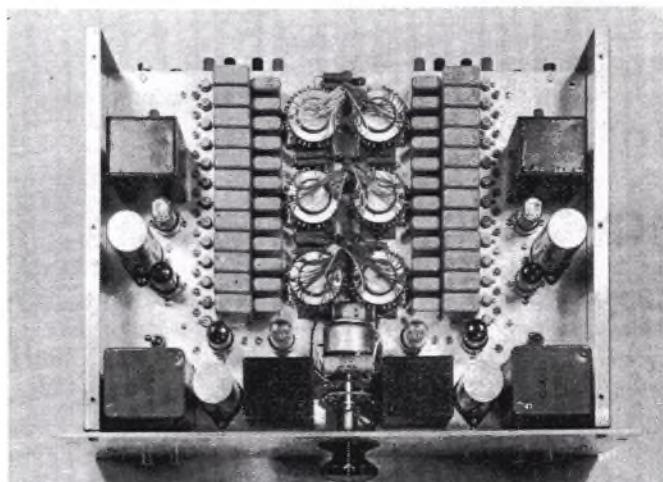


FIG. 8. Top view of the Dynamic Recording Correlator chassis.

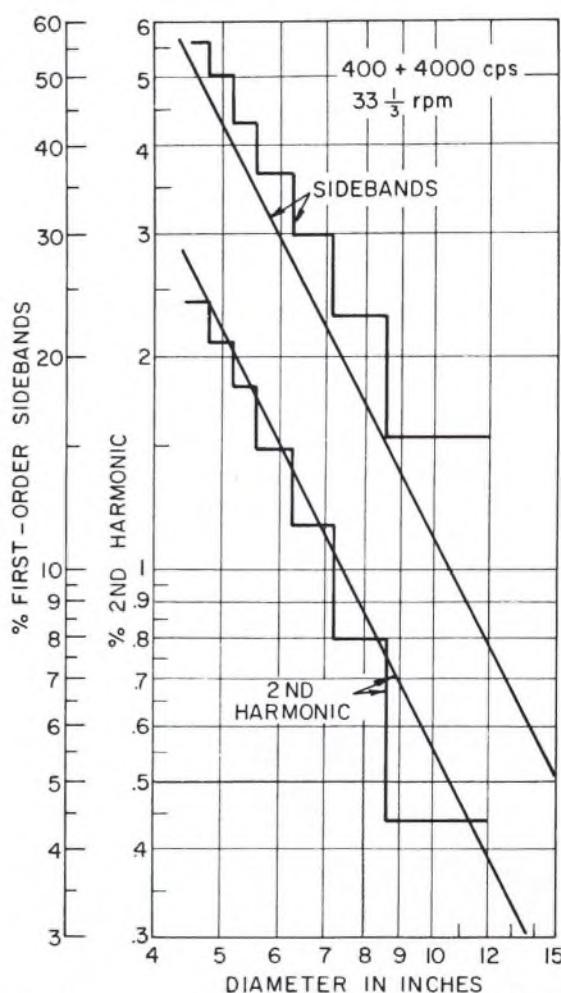


FIG. 9. Straight-line plots: Theoretical % first-order sideband and second harmonic components of tracing distortion. Step-shaped plots: Electrically-simulated tracing distortion in output of Correlator for each of 7 positions of diameter switch.

as are required. A topside view of the Correlator chassis is shown in Fig. 8. The Correlator is a unity-gain device operating from and into a 600-ohm program buss, and is normally inserted in the recording channel immediately preceding the RIAA recording equalization and the power amplifiers driving the recorder.

It was noted above that the total delay time required in the delay line is a function of the linear groove velocity and of the tip radius of the playback stylus. The Correlator has been designed to simulate playback with a 0.7-mil stylus tip radius, which is the most-commonly-used size in stereo pickups. Since the groove velocity varies as a function of the diameter, the delay time of the Correlator must also be adjusted as a function of the recording diameter. The required values of the delay time, δt , corresponding to diameters of 4.5 and 11.5 in. on a $33\frac{1}{3}$ rpm record are 0.493×10^{-3} and 0.126×10^{-3} sec, respectively. The Correlator is provided with seven different values of delay time which are selected by a stepping switch that closes a succession of multi-pole relays. The stepping switch may be operated either manually or automatically by means of contacts mounted on the

carriage of the recording lathe. The seven delay times were chosen to yield simulation of the playback process close to theoretical values for the following ranges of diameters: 11.5-8.6 in., 8.6-7.2 in., 7.2-6.3 in., 6.3-5.6 in., 5.6-5.2 in., 5.2-4.8 in., 4.8-4.5 in.

The degree to which the Correlator actually simulates the playback process can be judged by comparing the distortion measured at the output of the Correlator with the tracing distortion calculated for an ideal playback system. One such comparison is shown in Fig. 9 for a $400 + 4000$ cps signal having a 4:1 velocity ratio and a peak velocity in one 45/45 channel of 6.22 cm/sec. The theoretical amplitudes of the

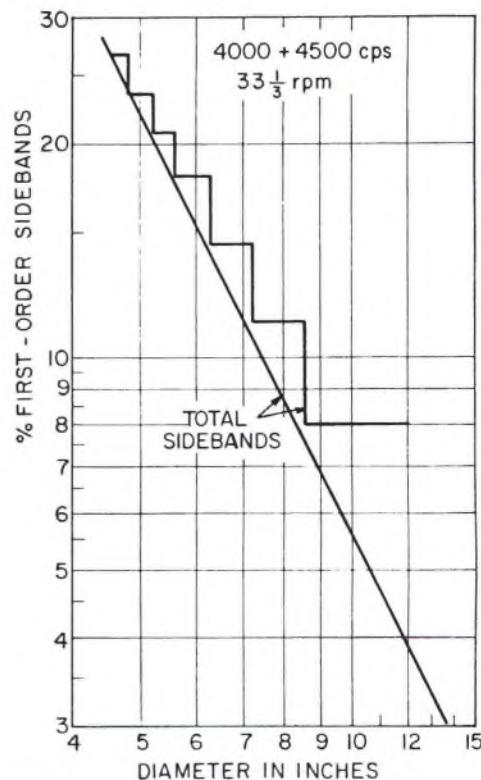


FIG. 10. Straight-line plot: Theoretical % first-order sideband components of tracing distortion. Step-shaped plot: Electrically-simulated tracing distortion in output of Correlator for each of 7 positions of diameter switch.

sidebands at 3600 and 4400 cps were calculated, added, and divided by the amplitude of the 4000 cps component to find the % first-order sidebands. This percentage was plotted as a function of groove diameter for a $33\frac{1}{3}$ rpm record to give the straight-line graph as shown on a log-log plot. The theoretical amplitude of the 800 cps component was calculated and divided by the amplitude of the 400 cps component to find the % second harmonic. This %, when plotted against groove diameter, also gives a straight line plot on a log-log chart. In the plots of Fig. 9 the data have been shown on the basis of a velocity-responsive pickup. The corresponding distortion products measured in the output of the Correlator for the 7 diameter-switch positions are shown in Fig. 9 as step-shaped plots.

A second comparison between the calculated distortion

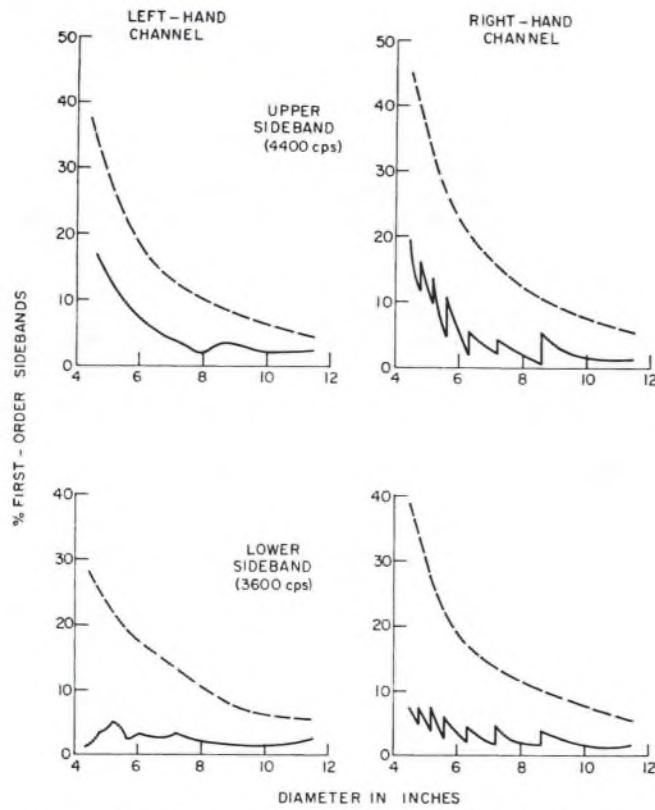


FIG. 11. Playback measurements of first-order sidebands of a 400 + 4000 cps signal recorded without (broken curves) and with (solid curves) the Dynamic Recording Correlator in the system.

and that produced by the Correlator is shown in Fig. 10. In this case the signal used consisted of 4000 + 4500 cps in a 1:1 displacement ratio. The amplitudes of the sidebands at 500 and 8500 cps on a displacement basis were added and divided by the amplitude of one of the original components to give the % first-order sidebands.

These results show that the Correlator gives a close, though not exact, simulation of theoretical tracing distortion. The discrepancies between theoretical and measured distortion can be attributed to the tolerances allowed in the selection of circuit components and in the adjustment of the stylus-shape voltages, to small residual non-linearities in the Correlator and associated equipment and instrumentation, and in the delay characteristics associated with a delay line having a finite number of sections. In principle the theoretical distortion could be simulated very exactly by a hyper-careful selection of circuit components and by use of a delay line of many sections. The 12-section line and the tolerances used for critical components represent a reasonable compromise between performance and complexity and cost. As will be seen, the existing residual discrepancies in simulation are small compared to distortions due to other variables in the record-playback processes.

RECORD-PLAYBACK TESTS

The ultimate test of the principle of electronic simulation of tracing distortion as embodied in the Dynamic Recording

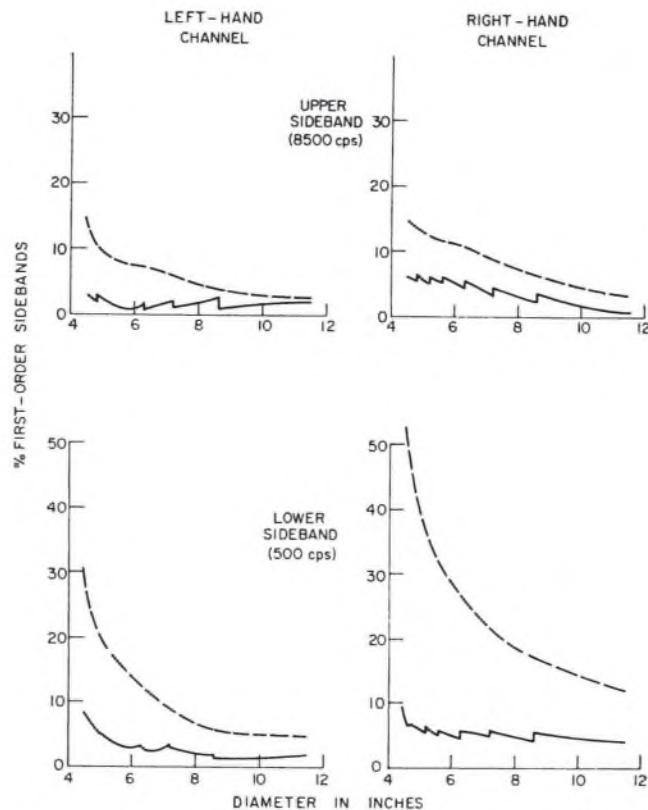


FIG. 12. Playback measurements of first-order sidebands of a 4000 + 4500 cps signal recorded without (broken curves) and with (solid curves) the Dynamic Recording Correlator in the system.

Correlator is, of course, its effectiveness in reducing distortion in actual record-playback operation. A number of types of test signal were employed in the record-playback evaluation of the Correlator. Typical results for two different intermodulation tests will be presented here. In each case the test signal was recorded as vertical modulation on a lacquer master, using a recorder adjusted to record with a 15° vertical angle in the modulation on the disk. In playback a pickup having a 15° vertical tracking angle was used. One signal consisted of 400 + 4000 cps in a 4:1 velocity ratio, recorded at a peak vertical velocity of 8.8 cm/sec. The other signal consisted of 4000 + 4500 cps, with each component recorded with a peak vertical displacement of 3.5×10^{-5} inch. Playback measurements were made on lacquer masters and also on vinyl pressings. No significant difference was noted in the two sets of measurements.

The test signals were recorded over most of the surface of 12-in. master disks, both with and without the Correlator in the recording system. When the Correlator was used its 7-position diameter switch was switched at the appropriate times. The playback measurements for the 400 + 4000 cps signal are shown in Fig. 11. A wave analyzer was used to measure the amplitude of the 4000 cps component and of the first-order sidebands in each channel. The dashed curves in Fig. 11 are plotted from the measurements of recordings made without the Correlator operating. The solid curves represent measured results when the Correlator was in op-

eration. The effect of advancing the diameter switch is evident in some cases. The data in this figure were measured and are presented on a velocity basis.

The playback measurements for the 4000 + 4500 cps signal are shown in Fig. 12. By means of a wave analyzer, the amplitudes of the components at 4000, 4500 and 500 and 8500 cps were measured. The dashed and solid curves refer, respectively, to tests without and with the Correlator. The data in this figure are presented on a displacement basis.

DISCUSSION

Several observations may be made with regard to the results shown in Figs. 11 and 12. The first and, at the moment, the most noteworthy is that the Dynamic Recording Correlator and the simulation techniques which it includes have reduced tracing distortion in playback very significantly for the two types of IM tests presented here. Other types of test signal have been used and have given comparable results. In particular, the quality of the reproduction of recorded music is noticeably improved at the innermost portion of the recorded area on a record when the Correlator is used during recording. For the tests reported here, the stylus-shape voltage characteristic was adjusted to represent a 0.7-mil stylus tip radius. The maximum reduction of distortion in playback is obtained when the tip radius of the pickup stylus has this same value. A lesser reduction of distortion will be obtained for either smaller or larger styli within practical dimensional limits.

A second observation, based on Figs. 11 and 12, is that the simulation technique was not always equally effective for both channels and both sidebands. A third observation, which is related to the second, is that the first-order sideband percentages measured when the Correlator was not used are almost always significantly higher than the theoretical values calculated and plotted in Figs. 9 and 10. These observations lead to the not-surprising conclusion that processes other than simple tracing distortion are also generating distortion products. A much more detailed experimental study will be required before the various processes responsible for the residual distortion can be identified and their contributions evaluated. However, a number of possible candidates can be listed in both the recording and the playback operations. Reference has already been made to the fact that in the present experiments the vertical tracking angle was adjusted to minimize distortion due to this cause. The type of distortion resulting from vertical-tracking-angle errors has been discussed in detail elsewhere.^{8,9} Further consideration discloses

an interaction between tracking-angle distortion and tracing distortion. This interaction is responsible for some residual distortion even when the tracking angle is properly adjusted and when the Correlator or equivalent techniques are used to remove the simple theoretical tracing distortion. The interaction becomes zero only when the vertical angle is zero, i.e., for strictly vertical motion, in both recording and playback. Preliminary calculations suggest that the residual distortion due to interaction could amount to a few percent of IM sidebands in the case of a 400 + 4000 cps signal under the test conditions used in the present study.

Any nonlinear element or mechanism in either recording or playback will produce intermodulation products which may combine with the products due to tracing distortion in various ways to yield erratic results such as are found in the measurements of Figs. 11 and 12. In recording, nonlinearities may exist in the recording amplifiers and in the cutter. Nonlinear mechanisms also result from the use of burnishing facets on the cutting stylus, from bending of the recording stylus and from springback of the groove walls of the lacquer master. The electrochemical duplicating and pressing processes may also affect the intermodulation products. In playback the distortion products may be affected by the lateral tracking angle, by elastic and plastic deformation of the groove walls by the stylus, by a radial skating force on the pickup and by nonlinearities in the pickup or the playback amplifiers.

The major inherent sources of distortion in stereodisk-record reproduction heretofore have resided in vertical tracking angle errors and in tracing effects. Techniques are now available for controlling both of these in large measure. Hence, as long as the basic slope and curvature overload limits in recordings are not exceeded, the distortion in the sound reproduced from stereodisk records may be made very low indeed, even with signals recorded at relatively high levels and at the innermost recorded area on the disk. The small amount of residual distortion is largely due to a combination of several second-order nonlinear phenomena as suggested in the preceding paragraph. Reducing the distortion due to these causes will require control of every element in the record-playback process within tolerance limits much closer than has been practiced heretofore.

⁸ J. G. Woodward and E. C. Fox, "A Study of Tracking-Angle Errors in Stereodisk Recording," *IEEE Trans. on Audio AU-11* (March-April, 1963).

⁹ B. B. Bauer, "The Vertical Tracking Angle Problem in Stereophonic Record Reproduction," *IEEE Trans. on Audio AU-11* (March-April, 1963).

Compensation for Tracing and Tracking Error

DUANE H. COOPER

University of Illinois, Urbana, Illinois

A method of compensation proposed by Professor F. V. Hunt to eliminate tracing distortion in the mastering of stereo disc recordings is shown to be equally effective against tracking distortion. The effectiveness is hardly diminished when used to compensate for the joint effects of both tracing and tracking, since approximately an order of magnitude of improvement seems to be offered, even in cases of severe error. The possibilities of practical application, to cure at one stroke the two major sources of distortion in stereo disc reproduction, are discussed.

INTRODUCTION

RECENTLY, in these pages, Professor Hunt cited an old trick by which tracing distortion might be removed from stereo disc reproduction.¹ One of the purposes of the present paper is to show that the same means will remove tracking distortion. It is a little disappointing to have to report that the compensation is exact only if one or the other of the two distortions is not present. However, the disappointment is only slight. It appears that even when both are present to a great degree, the application of Hunt's method will reduce the total distortion by roughly an order of magnitude.

Thus, we are presented with the opportunity of removing, almost completely, the two major sources of distortion present in stereo disc reproduction; it is gratifying that one measure will do for both.

COMPENSATION FOR TRACING ERROR

Tracing distortion² arises because the record groove, cut with a nearly knife-edged stylus, is, for reasons of slight wear, traced by a stylus with a spherical tip. In the tracing of mono records, the opposing groove walls carry comple-

mentary signals, producing, in the horizontal plane, a kind of push-pull geometry so that even-order distortion is prevented. In the stereo case, however, large amounts of second harmonic distortion can occur as well.

If the groove wall is thought to carry a plot of y , normal to the axis, versus x , measured along the quiescent groove axis, then the center of a sphere of radius a following the groove traces a curve given by x' and y' in the equations below. For reasons of compactness, these equations will be referred to as representing the transformation T_c :

$$\begin{aligned}x' &= x - a(dy/dx)[1 + (dy/dx)^2]^{-\frac{1}{2}}, \\y' &= y + a[1 + (dy/dx)^2]^{-\frac{1}{2}}.\end{aligned}\quad (1)$$

The distortion may be calculated from these equations to have a second harmonic content, when small, of

$$D(T_c) = 100 a / 4\rho_{min} \quad (2)$$

percent, in which ρ_{min} is the smallest radius of curvature of the recorded sinusoid. Intermodulation distortion has also been calculated.² Formula (2) is fairly accurate, even when $\rho_{min}=a$; the actual distortion in that case is nearer to 20% than to 25%. In the velocity mode, the distortion is twice that given here.

Professor Hunt's proposal is that a lacquer be cut in the usual manner, then traced by a stylus having the desired curvature; the signal so derived is to be used to drive the cutter for preparing the master lacquer, from which the

1. F. V. Hunt, "The Rational Design of Phonograph Pickups," *J. Audio Eng. Soc.* 10, 274 (1962).

2. M. S. Corrington, "Tracing Distortion in Phonograph Records," *RCA Rev.* 10, 241 (1949). Also with T. Murakami in the same journal for June, 1958.

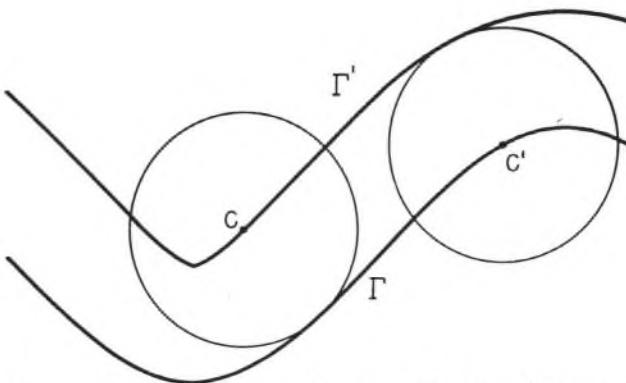


FIG. 1. Compensation for tracing error. The sinusoidal curve Γ , with peak slope 1.05, is followed by a circle of such radius as to match the sharpest curvature. The center at c traces the curve Γ' . Then Γ' is traced on its underside—equivalent to a polarity reversal of Γ' , followed by tracing, then followed by a polarity reversal—so that the center of the circle, at c' , retraces Γ exactly. The harmonic content of Γ' can be estimated in this case, by likening it to a full-wave rectified sinusoid, to be 20%.

pressings are to be derived. For the second lacquer the polarities are to be reversed from those cut into the first.

The validity of the compensation is most easily demonstrated by means of the geometrical construction of Fig. 1. The curve Γ , plotted by the cutter, is followed by a circle centered at c . The center c traces the curve Γ' , which is replotted by the cutter. Then Γ' is followed on its underside (polarity reversed) by the circle centered at c' . The center c' traces Γ back again, exactly.

In the diagram, the curvature of the circle just matches the sharpest curvature of the curve Γ . Thus, Γ' shows a sharp turning point. Sharper curvatures in Γ cause an abrupt change in tracing, and the compensation fails at those points.

COMPENSATION FOR TRACKING ERROR

Tracking distortion arises when the reproducing stylus does not move in the same direction, from the quiescent groove axis, as the direction the cutting stylus moves.³ Lacquer spring-back⁴ complicates the problem, so that the Westrex cutter, designed for a vertical motion 23 degrees from the true vertical,⁵ produces lacquers that would have been expected from a 0-degree cutter. Most high-quality magnetic cartridges are designed to track at 23 degrees from the vertical. European pressings require a backward tracking angle of some 16 degrees!

The equations, representing the tracking transformation, T_k (*tracing* is spelled with a "c," *tracking* with an additional "k"), may be given in one recursive formula:

$$y' = [\cos\psi/\cos(\psi+\phi)] f(x' + y' \sin\phi/\cos\psi). \quad (3)$$

Here, the original wave form is given by $y = f(x)$. The effective cutter angle is ψ , and the tracking error is ϕ . Analysis

based on Eq. (3) predicts the harmonic terms for simple sinusoids, and the intermodulation terms of both an amplitude and phase modulating character for double tones, that have been known since 1941.⁶ For the stereo case, phase cross-modulation has been demonstrated.⁷ In the velocity mode, the percent second harmonic is

$$D(T_k) = 100 \tan a \sin\phi/\cos(\psi+\phi), \quad (4)$$

when small. The peak slope of the sinusoid is $\tan a$. These and intermodulation terms can easily run to dozens of percents, with today's disagreement over the tracking angles.

The design of some cutters will not allow them to be tipped through large angles without extensive modification.⁸ Also, there is some question as to whether lacquer springback is wavelength independent.⁹ An automatic means of compensation, one that did not require precise measurement of complex dynamic phenomena, would be convenient.

That Professor Hunt's proposed compensation works for tracking error may also most easily be seen in a geometrical construction, shown in Fig. 2. It should be noted that the

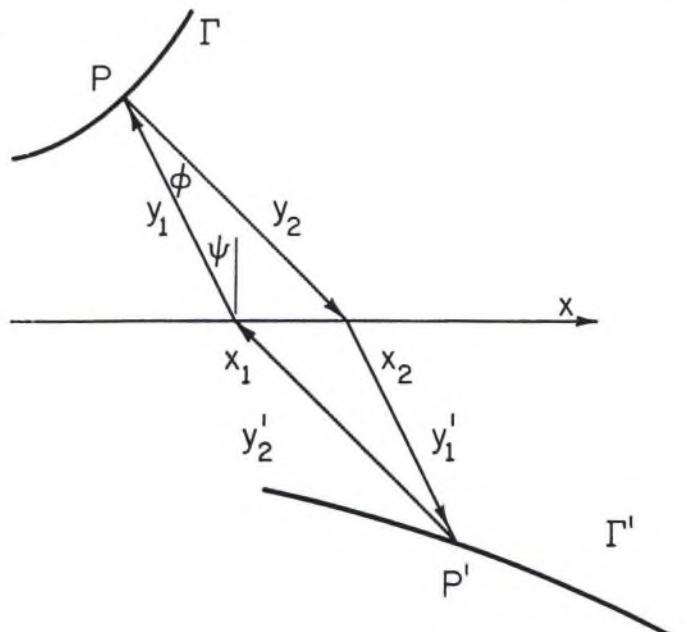


FIG. 2. Compensation for tracking error. The cutter plots P on Γ by moving from x_1 a distance y_1 in the direction inclined at angle ψ to the normal to the axis. The cartridge tracks with angular error ϕ , "reading" P as plotted a distance y_2 from x_2 . This reading is renormalized to remove the constant gain factor $\cos\psi/\cos(\psi + \phi)$, and drives the cutter to P' a distance of $y'_1 = -y_1$, but from x_2 . When Γ' is retracked with error ϕ , P' is found to be associated exactly with x_1 . The proof rests on the nature of a parallelogram.

6. H. G. Baerwald, "Analytic Treatment of Tracking Error and Notes on Optimal Pick-Up Design," *J. Soc. Motion Picture Engrs.* 37, 591 (1941).

7. D. H. Cooper, "Tracking Distortion as Phase Modulation," *IEEE Trans. on Audio AU-11*, 41 (1963).

8. C. R. Bastiaans, "Further Thoughts on Geometric Conditions in the Cutting and Playing of Stereo Discs," *J. Audio Eng. Soc.* 11, 6 (Jan., 1963).

9. C. Alexandrovich, "A Method of Measuring Vertical Angle," *Audio* 47, 6 (L) (Apr., 1963).

3. E. R. Madsen, "Vertical Tracking Angle—A Source of IM Distortion," *Audio* 46, 21 (Nov., 1962).

4. B. B. Bauer, "Vertical Tracking Improvements in Stereo Recording," *Audio* 47, 19 (Feb., 1963).

5. C. C. Davis and J. G. Frayne, "The Westrex Stereo Disc System," *Proc. IRE* 46, 1686 (1958).

distortion arises because of the displacement term in the argument of f in formula (3), for T_k . Thus, in Fig. 2, the cutter starts from the x -coordinate, x_1 , and moving a distance y_1 in the direction ψ , plots the point P on the curve Γ . That point is tracked by the cartridge as displaced a distance y_2 , but from x_2 . This "reading" is renormalized to remove the constant gain factor, $\cos\psi/\cos(\psi+\phi)$, and drives the cutter to replot at P' using $y'=-y_1$ but moving from x_2 . When Γ' is re-tracked, the point P' is found to be associated with x_1 , exactly. The proof rests on x_1Px_2P' being a parallelogram.

SIMULTANEOUS COMPENSATION

Abstractly, what Professor Hunt proposed is that the distortion in the waveform

$$f' = T_c f \quad (5)$$

could be removed, i.e., f recovered, by using the reflection transformation, R_y :

$$\begin{aligned} x' &= x, \\ y' &= -y, \end{aligned} \quad (6)$$

in combination with T_c , to generate the inverse, T_c^{-1} , of the tracing transformation. Here, these transformation symbols are interpreted as "doing" the transformation on everything written to the right. Thus, when more than one is to be applied, the right-most one operates first.

In this notation, if one follows Hunt, f is to be replaced by $R_y T_c f$ in Eq. (5). The geometrical proof of Fig. 1 shows that the result,

$$T_c R_y T_c f,$$

is the same as $-f$, or that

$$f = R_y T_c R_y T_c f,$$

or, that the identity transformation is

$$I = R_y T_c R_y T_c. \quad (7)$$

Multiplication on the right with T_c^{-1} in Eq. (7) shows that the representation of T_c^{-1} is

$$T_c^{-1} = R_y T_c R_y. \quad (8)$$

Multiplication on right and left with R_y in Eq. (8) shows, since R_y is self-inverse, that

$$T_c = R_y T_c^{-1} R_y. \quad (9)$$

Multiplication on only the right or the left with R_y in Eqs. (8) or (9) results in the formulae for manipulating the position of R_y :

$$\begin{aligned} R_y T_c &= T_c^{-1} R_y, \\ T_c R_y &= R_y T_c^{-1}. \end{aligned} \quad (10)$$

The argument for Fig. 2 shows that Eq. (7) is obeyed by T_k , after renormalization of the gain. Let T_{kn} denote the tracking transformation after the constant gain factor $\cos\psi/\cos(\psi+\phi)$ has been divided out. Then, it is T_{kn} that has the same properties given by Eqs. (7), (8), (9), (10), enu-

merated above for T_c . These properties are also shared by the rotation transformations T_r of the famous dihedral group.¹⁰ This fact is a consequence of each of these being dependent on its own parameter θ , such that

$$\begin{aligned} T(\theta_1) T(\theta_2) &= T(\theta_1 + \theta_2), \\ T(0) &= I = T(\theta) T(-\theta), \\ RT(\theta) R &= T(-\theta). \end{aligned} \quad (11)$$

Other properties of the dihedral group are not shared.

The tantalizing question, however, is whether so simple a means can be found to generate the inverse of $T_{kn} T_c$. The answer, as it develops, is that Professor Hunt's suggestion is very effective, although it does not produce an exact inverse. The procedure results in the overall transformation,

$$T_H = R_y T_{kn} T_c R_y T_{kn} T_c, \quad (12)$$

and the question turns upon the possibility of T_H being the identity transformation. It is easy to see that this question is not automatically answered on the basis of general principles. The formulae in Eq. (10), for example, can be invoked to clear T_H of reflections, so that

$$T_H = T_{kn}^{-1} T_c^{-1} T_{kn} T_c. \quad (13)$$

Thus, for T_H to be the identity, it is necessary and sufficient for T_{kn} and T_c to commute, that is, the commutator

$$T_{kn} T_c - T_c T_{kn}$$

must vanish. This property must turn on the details of the structure of T_c and T_{kn} . Either answer is possible; for example, the rotation transformation T_r can be shown to commute with T_c but not with T_{kn} .

A detailed analytic representation of Eq. (12) in terms of Eqs. (1) and (3) is elementary, but tedious and involved. While it does show that T_{kn} and T_c do not commute, that result can be seen more simply. Consider applying these transformations to a circle. That such a double-valued figure is not a waveform is irrelevant. Then, T_c carries circles into circles. Also, T_{kn} carries circles into ellipses. However, T_c carries ellipses into oval figures which can be, at best, only excellent approximations to the ellipse.

At the present time, it is feasible to exhibit numerical illustrations, derived by graphical methods, of the excellence of the approximation

$$T_H = T_{kn}^{-1} T_c^{-1} T_{kn} T_c \approx I. \quad (14)$$

One of these is shown in Fig. 3. The points marked a are the plotting points of a sinusoid. The sinusoid is traced by a circle whose curvature matches the sharpest curvature of the sinusoid. The tracing produces the curve at b . The tracing is tracked with a tracking error of 22° and used to drive a vertical cutter, plotting the curve with reversed polarity at c .

The curve at c shows the combined effect of both errors, $T_{kn} T_c$. The distortion is greater than produced by T_c alone, as at b . The curve at c is representative of the master which

10. H. Margenau and G. M. Murphy, *The Mathematics of Physics and Chemistry* (D. Van Nostrand Company, New York, 1943), p. 550.

embodies Hunt's compensation. The result of applying $T_{kn}T_c$ to it is yet to be shown. In so doing, it will be useful to pretend that the result will again be used to drive a vertical cutter with reversed polarities, to see how nearly the original sinusoid is recreated, i.e., by generating T_H as in Eq. (12).

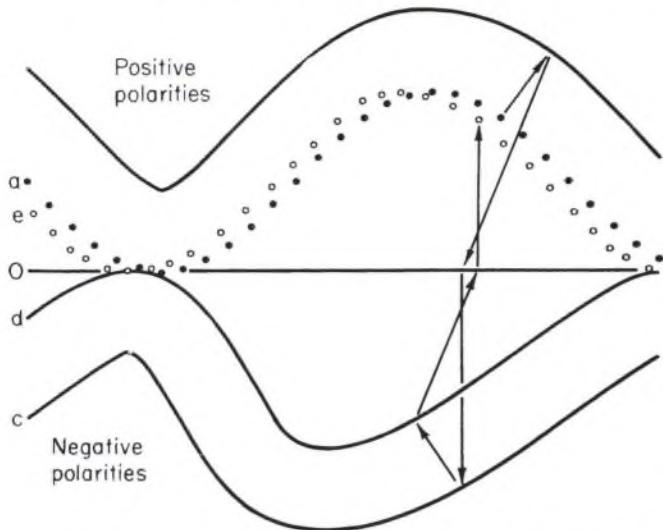


FIG. 3. Simultaneous compensation. The curve marked *a* is sinusoid with peak slope 1.05. The arrows show the sequence of tracing to *b* using a following circle whose curvature matches the sharpest curvature of *a*. Following the arrows, tracking with error 22° and replotting with reversed sign, carries the result to *c*. Then, retracing, retracking, and replotting carries the result to *e*, through *d*. The curve *e* is to be compared with *a*. The harmonic content of the curve *e* is judged to be a percent or so. Since the tracking distortion would give a second harmonic term of about 21%, and since that much distortion is already present in *b*, then, using the fact that the second harmonic term from tracking is in phase quadrature with that from tracing, the harmonic content of *c* is estimated at 30%.

The tracing construction based on *c* yields the curve *d*. This curve looks as if it involved only tracking distortion; there are no remarkably sharp turning points. Finally, the tracking construction leads to the points *e*. The closeness of these points to those of *a* invites close comparison.

There is some random-looking irregularity in the spacing of the points, which is probably a reflection of construction errors. In the presence of such irregularities, it is difficult to see any systematic deviations beyond a constant shift in

x . In a more accurate construction, changing ϕ to 30° produces a clearly noticeable increase in distortion. Figure 4 shows the compensated curve in which the distortion is estimated to be 6%, about a sixth of the uncompensated distortion. From this experience the residual distortion for Fig. 3 is judged to be a percent or so.

Until more numerical work can be done, it must suffice to say that the compensation results in distortion which may well be so small as to challenge experimental measurement, in less extreme cases.

PRACTICAL APPLICATION

There will be practical difficulties, of which an adequate survey will not be given here. Perhaps the greatest of these concerns the increased risk of introducing extraneous noises in the reproduction of the intermediate cutting. However, under laboratory conditions it should be possible to keep such noises to a much lower level than otherwise. Some distortions, such as aperture distortion, would be compounded. Also, the proposed cure does not work if the intermediate waveform is not traceable, either because the curvature becomes sharper than that of the reproducing stylus or because peak slopes become so steep as to nearly force a double-valued tracing.

The great charm of the proposal lies, as noted, in its ability to eliminate two sources of distortion at one blow. A second merit is that the whole burden of correction falls upon the reproducing cartridge, in tracing the intermediate cutting, so that the standard parameters such as stylus tip radius and vertical tracking angle can be changed over great limits without requiring changes in cutter design. Also, because it is subjected to lighter stresses, a precision cartridge is better able to bear the burden of meeting the performance specifications expected of a standard. Even lacquer spring-back and effects due to elastic deformation in the cutter itself are automatically compensated, so long as they are reproducible from one cutting to the next. There is one difficulty. Lacquer springback, which affects ψ , appears to be slightly wavelength dependent.¹⁰ This is involved in re-normalizing the "constant" gain factor $\cos\psi/\cos(\psi+\phi)$. Fortunately, errors in ψ affect the "constancy" of this gain to hardly more than second order, so that possibly the overall effect is third or fourth order. Otherwise, changes in such parameters have no effect and they require hardly any monitoring or allowing for, let alone precise measurement.

In placing the burden of standardization upon the cartridge used to play back the intermediate cutting, it is necessary to recognize the special requirements the design of such a cartridge must meet. First, it may be pointed out that its role as a laboratory calibration standard relaxes some limitations. Principally, these are the ones of cost, sensitivity, and durability of replaceable parts. The new limitations come from the requirements of precisely defined and calibrated geometric, dynamic, and electrical parameters, together with the choosing of the values of such parameters at the boundaries of the state of the art.

As an example, consider the demountable stylus-armature

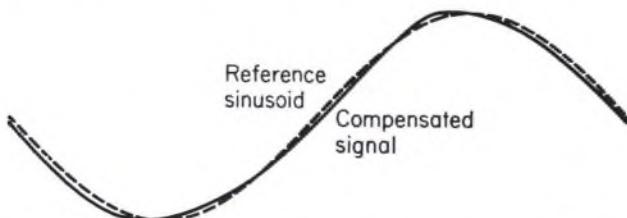


FIG. 4. Residual distortion. The solid curve shows the effect of trying to compensate for a 30° tracking error in the presence of the extreme tracking error of Fig. 1. The dashed curve is a reference sinusoid. The portion of sharpened curvature near the maximum is real and makes its appearance so noticeably when increasing ϕ from 22° to 30° . The tracking distortion alone would be 32%. Quadrature combination with a 20% tracking distortion results in a 38% figure for the uncompensated distortion. After compensation, some 6% is estimated.

assembly. One would want to strive for a precisely defined pivot, without translational softness.¹¹ The calibrated compliance should be as great, and the calibrated damping as slight, as good design will permit. Perhaps more important is the effective stylus tip mass, to be chosen as low as the requirement of armature and cantilever stiffness will allow. Surrendering the requirements of durability, cost, and sensitivity may make possible great advances in low tip mass. The use of sapphire suggests itself for the tip, because it may be more precisely shaped and polished than diamond, and thus may be more adaptable to low mass designs. This one parameter is important because the greater compliance of the acetate lacquer makes tip-mass distortion likely to be more significant than with vinyl, and because such distortion can compromise the accuracy of tracing compensation.¹² Calibrated effective stylus tip masses near or below a tenth of a milligram are to be desired.

As such advances appear, the precision with which compensation may be performed will be advanced. The present state of the art probably allows significant precision in compensation. These advances will ultimately affect the design of cartridges offered for home use, although the home requirements are less severe than those of the cutting laboratory. A tracing-compensated groove, for example, requires the stiffer vinyl to supply smaller accelerations to the stylus tip than does the uncompensated groove. It happens that a groove compensated for a half-mil stylus radius may be expected to show an improvement, even when played with a one-mil stylus, over the uncompensated one. A standard parameter, such as stylus radius, may thus be reconsidered as the state of the art advances.

For the standardization of tracking angles, it is natural to want to see the past five year's investment in stereo discs protected. That cause seems almost lost, because of the difficulty of designing high-quality cartridges for zero or negative vertical tracking angles. Some relief will be available through the reprocessing of master tapes of favorite performances in accordance with the new disc standards. For some unknown number of stereo discs relief is little needed, since they were already cut with extremely conservative vertical levels, so much so as to sound quite decent in the presence of tracking and tracing error. In any case, the present scheme of compensation could make such considerations a thing of the past.

Since compensation requires no tipping of cutter mechanisms, there now seems little reason to avoid rather substantial positive values for the standard for the vertical tracking angle. Even 23° would seem almost reasonable, especially in view of the successful cartridge designs for that angle.

Of course, a conservative position would be that compensation be not required to work too large a change in angle from that which the cutter-lacquer interaction naturally produces. Eventually, we will understand all of the variable dynamic, elastic, and plastic factors so well that it will be possible to stabilize them fully; little compensation will then be needed, since the requisite cutter designs will be at hand. Even then, tracing compensation may be undertaken by this method with the assurance that not only any remaining tracking error will not compromise the tracing compensation, but that such error will be removed as well. At the present, however, the proposed compensation can accomplish so much that there seems to be little need to wait for further refinements in the theory.

Note Added in Proof: The expression for T_H , given above, and the results derived from it, are exact only if a cutter capable of giving a truly vertical cut is to be used, possibly such a one as the *Westrex 3c*. A paper is in preparation treating the cases of more general inclinations, the corresponding compensation, and the calculation of the residual distortion. The conclusions given here, however, accurately represent those for the more general cases.

THE AUTHOR



Duane H. Cooper was born in Gibson City, Illinois on August 21, 1923. He received the B.S. degree in 1950 and the Ph.D. degree in 1955, both from the California Institute of Technology, Pasadena.

Since 1954 he has been with the Coordinated Science Laboratory (then the Control Systems Laboratory) of the University of Illinois, Urbana, where he is Research Associate Professor of Physics and Electrical Engineering. In recent years his principal interests have been statistical detection, information and decision theory, with applications to pattern recognition, and the applications of digital computing systems to instrumentation and simulation of physical systems by Monte Carlo methods.

Dr. Cooper is a member of the American Physical Society.

11. B. B. Bauer, *op. cit.*

12. J. Walton, "Stylus Mass and Distortion," *Wireless World* 69, 171 (Apr., 1963).

Integrated Treatment of Tracing and Tracking Error *†

DUANE H. COOPER

University of Illinois, Urbana, Illinois

It is shown that tracing and tracking error are twin sources of distortion, each describable by the skew transformation. Means for exact Fourier analyses, applicable to both, are given. Various representations are discussed for the light they shed on compensation principles. A demonstration is given of the degree to which the present vertical modulation handicap, due to the joint action of both errors, may be removed.

INTRODUCTION

HILL-AND-DALE recording, or vertical groove modulation, has always had to contend with the problem of tracing error. This was recognized as early as 1932.¹ The problem was given its present name some five years later in the first of a series of papers devoted to its analysis.²⁻⁶ The revival of vertical modulation, to carry the separation or difference signal between the two channels of stereo disc recording, again focussed attention on this classic problem, although briefly.^{7,8}

Almost by chance, another old problem which had not been thought to be of especial significance for vertical modulation, that of tracking error,⁹ has also come into prominence with the advent of stereo disc recording. In the space of one year, it has been studied intensively.¹⁰⁻¹⁵ Whether tracing or tracking error is the more serious has been debated.^{16,17} The peculiar circumstances which brought these two classic problems into juxtaposition undoubtedly contributed to their being recognized as twins.¹⁸

That tracing and tracking error are twins derives from their common mathematical parentage, the skew transformation.¹³ It is fortunate that they may thus be accorded an integrated treatment in analysis, representation, and calculation of limits on the vertical recording level imposed by their joint action in generating distortion, for it turns out that these two errors are comparable in significance for current practice. The integrated treatment extends to possible compensation processes¹⁹ that promise to lay these two classical problems finally to rest.

* This work was done in part as consultant to Consumers Union, Mount Vernon, New York.

† Presented October 18, 1963, at the Fifteenth Annual Convention of the Audio Engineering Society, New York.

THE SKEW TRACING TRANSFORMATION

That tracing error may be described as a curvilinear skew transformation is shown with the help of Fig. 1. The lower part of that figure shows the profile S of a stylus tip engaging the waveform Y which modulates the wall of a phonograph groove. It is convenient to think of the record as being stationary and the stylus as being moved to trace

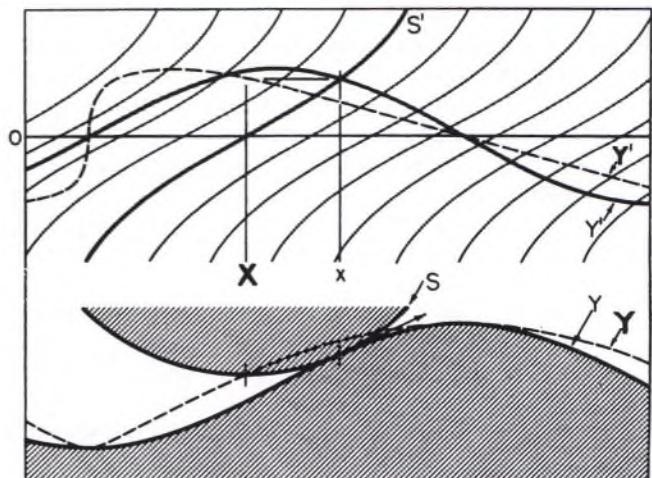


FIG. 1. The skew tracing transformation. The lower part of the figure shows the geometrical relationship of the stylus profile S to the modulated groove wall Y . The tangential contact is at x . The arrow shows the tangential direction the stylus would move, relative to the groove wall, in tracing the curve Y . Following a particular point on S , the vertex at x develops the tracing Y . In the upper part of the figure, the slopes of these curves, S' , V' , Y' , are plotted. The matching of slopes at x for S and Y is shown by the intersection of S' and Y' at x . Also, the ordinate of the Y' curve at x matches, because of the tangential motion of S on Y . Thus, the plotting of Y' is shown to invoke the curvilinear skew coordinates.

the waveform. The stylus contacts the groove wall at x and is constrained thereby to move in the tangential direction indicated by the arrow. The stylus moves as a rigid body so that every point fixed in it traces the same curve as any other. The vertex of the stylus profile at x is chosen as tracing the distorted curve Y .

The slopes of these curves are plotted in the upper part of Fig. 1. The curves S and Y have the same slope at x , so that the curves S' and Y' intersect there. The curve S has zero slope at x , so that S' has its intercept with the x -axis there. Because S is constrained to move tangentially, the curve Y has the same slope at x as Y has at x . Thus Y' has the same ordinate at x as Y has at x . The plotting of Y' may be described as follows: from the intersection of S' with Y' , move in the negative x -direction to a point just above the intersection of S' with the x -axis, and there plot a point of Y' .

These same words describe plotting the skew transformation,¹³ except that reference to S' would be replaced by reference to a tipped axis of ordinates. This observation is emphasized by showing in Fig. 1 a parallel family of S' -curves. These curves have odd symmetry if the stylus has even symmetry. For a parabolic stylus profile, the S' -curves are straight lines and the transformation is the ordinary linear skew transformation. Even for the circular profile shown, the departure from linearity is slight for moderate peak slopes in Y .

It is evident from Fig. 1 that the phase perturbation, $x-x$, is given by a function inverse to that describing S' . If this inverse function be called $cg(y')$, where y' is the ordinate of the Y' -curve, and c is the vertex radius of curvature of the stylus profile, then the transformation equations, T_c , are

$$x = x - cg(y'), \quad y' = y'. \quad (1)$$

The equations usually given⁸ are for the integral of Eq. (1), with

$$g = y'/[1 + (y')^2]^{\frac{1}{2}},$$

as for the circular stylus profile. For the parabolic tip, Eq. (1) becomes simply

$$x = x - cy', \quad y' = y', \quad (2)$$

and the equation for y is

$$y = y - (c/2)(y')^2.$$

The appearance of a non-linear amplitude perturbation, upon conversion to an amplitude basis, betrays the essential non-linearity of the phase-modulation process, however linear the relationship of the phase perturbation to the signal variables. In a later section dealing with sequences of skew transformations, some linear only on a slope basis, some linear only on an amplitude basis, it will be necessary to abandon linear methods of analysis.

THE SKEW TRACKING TRANSFORMATION

The distortion arising from vertical tracking error has been abundantly discussed.¹⁰⁻¹⁸ For the present purposes,

however, it will prove useful to discuss it in normalized factored form. The gain is to be normalized, and the skew transformation of cutting is to be separated from that of tracking.

If a cutter, working with an effective skew angle ψ , measured from the vertical, "plots" y as a function of x in its skew coordinates, the resulting curve in rectangular coordinates x, y , is given by

$$x = x - y \sin \psi, \quad y = y \cos \psi.$$

If the cutter, instead of plotting $y = f(x)$, were driven a little harder to plot $y = f(x)/\cos \psi$ in its skew coordinates, then the resulting maximum values in rectangular coordinates would be those of f , i.e., there would be recorded

$$x = x - f(x) \tan \psi, \quad y = f(x).$$

Thus the desired normalization is achieved.

The parameter of skewness $\tan \psi$ is to be written simply as k , and the skew cutter transformation as T_k :

$$x = x - ky, \quad y = y. \quad (3)$$

A similar description of the skew tracking transformation, upon reproduction, may be given. However, that transformation has the reverse sense from that for the cutter, so that if the parameter of skewness be called k , the transformation is to be called T_{-k} :

$$x = x + ky, \quad y = y. \quad (4)$$

The effect of more than one of these transformations being applied in sequence will be discussed in a later section.

FOURIER ANALYSIS

Exact Fourier analyses representing the application of the skew transformation to single tones and pairs of tones, as well as to the circumstances of cross-modulation, have already been presented in an earlier paper.¹⁸ The present method is simpler and gives the same results. It is applicable to the analysis of both tracking and tracing error, provided the latter can be represented by a linear skew transformation.

Insertion of the equations

$$x = x - ky, \quad y = y, \quad (5)$$

into the Fourier analysis integral for y' ,

$$H = (1/\lambda) \int e^{-j\omega x} y'(x) dx, \quad (6)$$

leaves the result

$$H = (1/\lambda) \int e^{-j\omega(x-ky)} y' dx, \quad (7)$$

since the exponent substitution expresses the first of Eqs. (5), and indefinite integration in Eqs. (6) and (7), for $\omega = 0$, reproduces the second. The variable y is to be periodic with angular frequency $2\pi/\lambda$, and ω is to be a non-zero whole multiple of that frequency. The definite integral is to span one period, λ . These observations allow integration by parts, transforming Eq. (7) to

$$H = (1/\lambda k) \int e^{-j\omega(x-ky)} dx, \quad (8)$$

which serves as the basis for the Fourier analysis of y' , and hence of y for a variety of finite linear combinations of simple oscillations for y .

As an example, consider

$$y = A \sin \omega_0 x.$$

The integral is well known,²⁰ and the resulting Fourier cosine coefficients are given in terms of Bessel coefficients, J_n :

$$2H_n = (2/k) J_n(nkA\omega_0).$$

The corresponding coefficients for the amplitude waveform are*

$$(2H_n/n\omega) = A(2/n\beta) J_n(n\beta), \quad (9)$$

in which β is the phase automodulation index, $kA\omega_0$. The analysis for the tracing transformation is analogous, and gives results in exact agreement with the expansions of Corrington,⁶ except for second-order peak-slope dependent terms, which are characteristic of the spherical stylus.

REPRESENTATIONS

It is useful to know as many representations of the skew transformation as possible, especially if compensation schemes are to be devised and analyzed. For example, the linear skew transformation may be represented by the matrix

$$T_k = \begin{pmatrix} 1 & -k \\ 0 & 1 \end{pmatrix}, \quad (10)$$

displaying the coefficients of the linear equations—Eq. (3). Then, the application of a sequence of such transformation may be represented with the help of the familiar rules of matrix multiplication.²¹

In writing down sequences of these transformations, the convention will be followed that the operation proceeds to the right. Thus, the application of T_k to $y = f(x)$ is written

$$y(x) = T_k f(x), \quad (11)$$

as representing Eq. (3). Often, explicit mention of the operand will be suppressed. For a sequence of operations, the rule of Eq. (11) implies that the operation standing for the right-most symbol will be applied first. Thus

$$T_{-k} T_c T_k$$

stands for the cutter transformation T_k operating first, followed by the tracing transformation, and finally followed

* *Historical Note:* Baerwald (Ref. 9) obtained this result in 1941 by a complicated method of analysis. He did not believe that the corresponding result for linear combinations of tones could be given exactly. Eugenio Fubini Ghiron also obtained this result in 1935 ("Anomalia nella Propagazione di Onde Acustiche di Grande Ampiezza," *Alta Frequenza* 4, 530). His style of analysis was the cumbersome one reinvented by the present author for a recent paper (Ref. 18), but he gives results only for the single-tone case. This 1935 paper was rediscovered in 1960 by W. Keck and R. T. Beyer ("Frequency Spectrum of Finite Amplitude Ultrasonic Waves in Liquids," *Physics of Fluids* 3, 346).

by the tracking transformation, T_{-k} . The evaluation of this combination will be given in the next section.

The combination

$$T_{-k} T_k = T_{k-k}$$

is evaluated by performing the indicated matrix multiplication:

$$\begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & -k \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & k-k \\ 0 & 1 \end{pmatrix},$$

whereby it is seen that, if the cutter and tracking angles agree, one has the identity transformation:

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad (12)$$

so that T_{-k} is the inverse of T_k :

$$T_{-k} = T_k^{-1}, \quad (13)$$

since their product is the identity transformation.

There will be occasion to use the reflection transformation for y ,

$$R_y = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad (14)$$

and for x ,

$$R_x = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}. \quad (15)$$

These transformations are the same when written for the application to slopes. With their help, an inverse to the tracing transformation may be given. Using R_y , one has

$$R_y T_c R_y = T_{-c} = T_c^{-1}, \quad (16)$$

as may be seen by performing the matrix multiplication,

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & -c \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 1 & c \\ 0 & 1 \end{pmatrix}.$$

As Professor Hunt has pointed out,⁵ such a procedure was proposed by MacNair for the compensation of tracing error.²² One would make an intermediate cutting with reversed polarities, then trace it with the intended stylus, and use the resulting signal to make a second cutting, again with reversed polarities. The tracing error in the reproduction of the second cutting is fully compensated, in the absence of any other errors, since the overall transformation would be

$$I = T_c R_y T_c R_y.$$

Equation (16) holds also for non-parabolic stylus profiles, provided they show even symmetry.¹⁸

Obviously, Eq. (16) holds for T_k replacing T_c . It also holds for R_x replacing R_y . This agrees with Bauer's observation¹⁴ that reversing the direction of playback reverses the sense of the cutting angle. Of course, the program would have had to have been recorded in reverse order as well. Thus, cutters with negative angles can be converted to positive ones and vice versa.

Possibly the most flexible representation is that of Fig. 2. It is an operational representation invoking elements of an analog computing system. Automodulation of the phase is accomplished by varying the length of a delay line, or by varying the delay constants of an otherwise fixed line. Studies are being made of such lines.^{23,24}

The amplifiers shown may embody differentiation or integration for conversion to and from the velocity mode, for

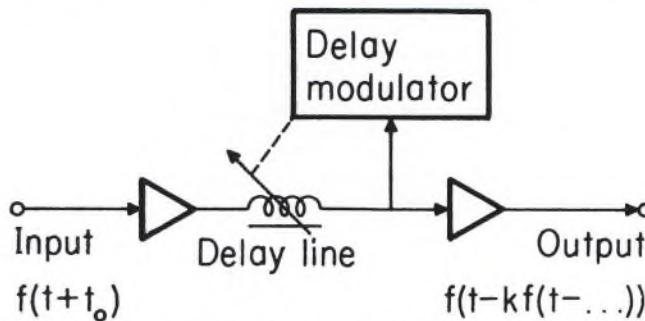


FIG. 2. Analog computer for the skew transformation. The requisite phase automodulation is achieved by modulating the delay constant or the length of a delay line. For adjustment of sign and magnitude of the modulation, the amplifiers may (or may not) be inverting and of adjustable gain. They may involve integration, differentiation, or neither, depending upon whether the phase modulation is to apply to amplitude or velocity. A non-linearity of odd symmetry in the modulator would allow representation of tracing by a stylus of even symmetry.

representing tracing. They may embody gain adjustments for controlling the phase-modulation constant. Again, they may embody polarity reversal for adjusting the sign of the modulation constant in accordance with Eq. (16). Finally, the delay modulator could exhibit a non-linear transfer function of odd symmetry to simulate the effect of a non-parabolic stylus.

The function of continued argument,

$$y = f[t - kf(t - \dots)], \quad (17)$$

shown in Fig. 2, may be deduced by recursive interpretation of the equations for T_{-k} . It constitutes yet another representation of the skew transformation. Truncation of the recursion to

$$y \approx f[t - kf(t)] \quad (18)$$

suggests approximate analysis by the method of phase perturbation.¹³

COMBINED TRACING AND TRACKING ERROR

The matrix multiplication method of calculating combined transformations fails when the resulting equations are nonlinear. Unfortunately this is the case even for $T_c T_k$. In such circumstances substitution methods may be used. The equations for $T_{-k} T_c T_k$, derived by substitution, are

$$\begin{aligned} x &= x - (k - k)y - cy' / (1 - ky') - (kc/2)(y')^2 / (1 - ky')^2, \\ y &= y - (c/2)(y')^2 / (1 - ky')^2, \\ y' &= y' / [1 - (k - k)y']. \end{aligned} \quad (19)$$

These equations represent the overall effect of recording

with skewness k and playback with a parabolic stylus of radius c and cartridge skewness k .

It is seen that even when $k = k$, the effects of skew cutting do not entirely vanish, as they would if T_c did not intervene. Thus it is observed that

$$T_{-k} T_c T_k \text{ not equal to } T_{-k} T_k T_c = T_c,$$

or, in other words, that T_k and T_c do not commute.

For the purposes of making approximate analyses, adequate for small distortions, one can justify the linearization of Eq. (19) to the approximate equations,

$$\begin{aligned} x &\approx x - (k - k)y - cy', \\ y &\approx y - (c/2)(y')^2, \end{aligned}$$

in which it is seen that the dominant tracing and tracking terms are in phase quadrature. This fact has been observed also by Woodward and Fox,¹⁵ and the observation extends to the distortion terms whether measured by harmonic content or by intermodulation indices. Then the total distortion arising from both kinematic errors may be obtained by a root-sum-of-squares combination of the distortion figures from each;¹⁸ conversely, if limits are placed on the maximum allowable distortion, one may derive limits for the maximum recording velocities imposed by both tracing and tracking error.

As an example, these limits have been calculated for a maximum harmonic content of 2% or a maximum intermodulation of 10%, whichever provided the more severe limitation. For intermodulation, the upper frequency was chosen to be 5,000 cps. A groove speed of 10 ips was chosen, and the stylus was taken to have a half-mil radius.

These limitations are shown plotted in Fig. 3. It should

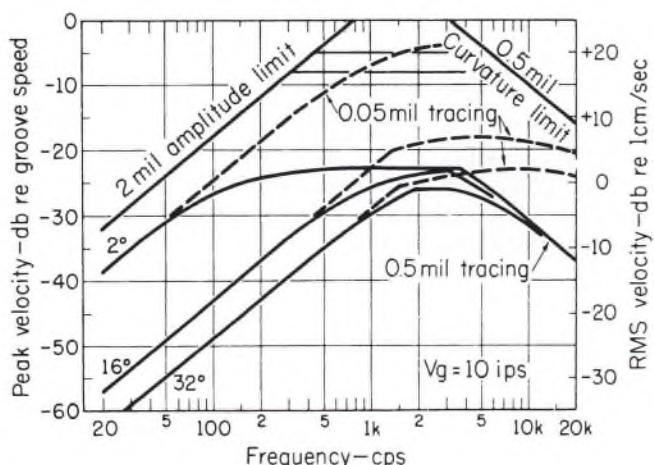


FIG. 3. Limitations on vertical modulation. The curves show sharp corners. At frequencies below these corners, ten-percent intermodulation distortion sets the limitation, and, for the frequencies above, two-percent harmonic distortion sets the limitation. These distortion figures reflect the joint effect of tracing and tracking error. Reducing the tracking error alone provides the relief shown. The dashed curves show the relief provided by reducing the effects of tracing error by an order of magnitude. When both errors are small, the relief is substantial. The horizontal lines at 5 db and 8 db below the groove speed show the ten-percent intermodulation limits left by the MacNair-Hunt compensation scheme for two cutter angles 0° and +15°, respectively.

be pointed out that they apply to the vertical modulation. The sharp corner in the curves shows the boundary between limitation by intermodulation at the lower frequencies, and limitation by harmonic content in the higher frequencies. The limitations are quite severe, although reduction of tracking error to only 2° does provide some relief in the lower frequencies.

Also, if some means could be found to reduce the tracing error by an order of magnitude, some relief would be obtained in the higher frequencies. However, correcting only one or the other of these errors would still leave the vertical modulation severely handicapped in dynamic range. The uppermost dashed curve is for a tracking error of 2° and a tracing error of 0.05 mil, representing a substantial reduction in both errors simultaneously. It allows a peak vertical modulation substantially as great as is customary for the lateral.

The means by which such high modulation levels are to be obtained must rest with one or another of the compensation schemes to be discussed shortly. This is not to say that such modulation levels have not been attained without the help of compensation. They have, and the resulting products have often been hailed as presenting challenges to reproducing equipment. It would be more accurate to say that such products presented vain, if unwitting, challenges to kinematics.

Even with compensation, the use of high modulation levels would still present some genuine challenge to the reproducing apparatus. The challenge would be to the stylus admittance, especially that part governed by effective stylus mass,²⁵ a parameter more properly within the province of cartridge design. A part of the challenge would also be presented to the stiffness of the record material.

COMPENSATION PRINCIPLES

If the compensation is to take account of the presence of both tracing and tracking error, it must be arranged, for maximum accuracy, according to the rules of non-commutative algebra. For example, if a truly vertical cut is to be made, then reproduction, which invokes

$$T_{-k} T_c,$$

is the only error. Then, the compensation to precede cutting must invoke the inverses of these in reverse order:

$$T_{-c} T_k.$$

These inverses may be generated by some realization of Fig. 2, for example. The overall transformation would be

$$(T_{-k} T_c)(T_{-c} T_k), \quad (20)$$

which is the identity.

The more general case is that of the skew cutter. Cutting and reproducing gives the overall transformation

$$(T_{-k} T_c) T_k, \quad (21)$$

shown in Eq. (19). The inverse to this is

$$T_{-k} T_{-c} T_k,$$

so that if compensation were to precede cutting, the overall transformation is to be

$$(T_{-k} T_c) T_k (T_{-k} T_{-c} T_k), \quad (22)$$

which, again, is the identity. It may be seen that no particular advantage attaches to the tipped cutter in these schemes. Its skew transformation merely has to be undone before-hand, in the compensation.

A compensation scheme which does not obey these rules would be the one of MacNair and Hunt. This scheme was originally proposed to compensate for tracing error, but it would be just as effective for removing tracking error, if either of these errors occurred singly, as has been seen above. The overall transformation for the general case would be

$$T_{-k} T_c (T_k R_y) T_{-k} T_c (T_k R_y), \quad (23)$$

which is

$$T_{-k} T_c T_k T_k T_{-c} T_{-k}, \quad (24)$$

and which is not the identity, in general.

As an example of a simple scheme¹⁹ violating the present rules, it is interesting to see how nearly Eq. (24) approximates the identity. The transformation equations representing Eq. (24) may be shown to be

$$\begin{aligned} x &= x - (3c/2)(k+k)(y')^2(N/D), \\ y' &= y', \end{aligned} \quad (25)$$

in which N/D is nearly unity, since

$$\begin{aligned} N &= 1 - (2/3)(k-k)y' - (1/3)kk(y')^2, \\ D &= (1-ky')^2(1+ky')^2. \end{aligned} \quad (26)$$

The resulting distortion would be small, the modulation index being essentially the product of the tracing and tracking indices that would have occurred.

In Fig. 3, velocity limitations set by the residual distortion in the MacNair-Hunt scheme are shown as the horizontal lines at 5 and 8 db below groove speed. It is the 10% intermodulation limit that governs. The 8 db limitation is for a skew cutter whose angle matches that of the cartridge at 15° . The 5 db limitation is for the same cartridge but a 0° cutter.

Finally, it may be noted from Eqs. (24) or (25) that, for a cutter angle precisely opposite to that of the cartridge, the MacNair-Hunt scheme provides exact compensation for the joint action of both kinematic errors.

REFERENCES

1. H. A. Frederick, "Vertical Sound Records," *J. Soc. Motion Picture Engrs.* 18, 141 (1932).
2. M. J. Di Toro, "Distortion in the Reproduction of Hill-and-Dale Recordings," *J. Soc. Motion Picture Engrs.* 29, 493 (1937).
3. J. A. Pierce and F. V. Hunt, "Distortion in Sound Reproduction from Phonograph Records," *J. Soc. Motion Picture Engrs.* 31, 157 (1938).
4. J. A. Pierce and F. V. Hunt, "On Distortion in Sound Repro-

- duction from Phonograph Records," *J. Acoust. Soc. Am.* 10, 14 (July, 1938).
5. W. D. Lewis and F. V. Hunt, "A Theory of Tracing Distortion in Sound Reproduction from Phonograph Records," *J. Acoust. Soc. Am.* 12, 348 (January, 1941).
 6. M. S. Corrington, "Tracing Distortion in Phonograph Records," *RCA Rev.* 10, 241 (1949).
 7. C. A. Davis and J. G. Frayne, "The Westrex Stereo Disc System," *Proc. IRE* 46, 1686 (1958).
 8. M. S. Corrington and T. Murakami, "Tracing Distortion in Stereophonic Records," *RCA Rev.* 19, 216 (1958).
 9. H. G. Baerwald, "Analytic Treatment of Tracking Error and Notes on Pick-up Design," *J. Soc. Motion Picture Engrs.* 37, 591 (1941).
 10. E. R. Madsen, "Vertical Tracking Angle—A Source of IM Distortion," *Audio* 46, 21 (November, 1962).
 11. C. R. Bastiaans, "Further Thoughts on Geometric Conditions in the Cutting and Playing of Stereo Discs," *J. Audio Eng. Soc.* 11, 6 (1963).
 12. B. B. Bauer, "Vertical Tracking Improvements in Stereo Recording," *Audio* 47, 19 (February, 1963).
 13. Duane H. Cooper, "Tracking Distortion as Phase Modulation," *IEEE Trans. on Audio AU-11*, 41 (March-April 1963).
 14. Benjamin B. Bauer, "The Vertical Tracking Angle Problem in Stereophonic Record Reproduction," *IEEE Trans. on Audio AU-11*, 47 (1963).
 15. J. G. Woodward and E. C. Fox, "A Study of Tracking Angle Errors in Stereodisk Recording," *IEEE Trans. on Audio AU-11*, 56 (1963).
 16. J. v.d. Steen, "On Geometric Conditions in the Cutting and Playing of Stereo Discs" (Letter), *J. Audio Engr. Soc.* 11, 280 (1963).
 17. Duane H. Cooper, "Further Comments on Geometric Conditions in the Cutting of Stereo Records (Letter)," *J. Audio Engr. Soc.* 11, 280 (1963).
 18. Duane H. Cooper, "Unified Analysis of Tracing and Tracking Error," *IEEE Trans. on Audio AU-11* (July-August, 1963).
 19. Duane H. Cooper, "Compensation for Tracing and Tracking Error," *J. Audio Eng. Soc.* 11, 318 (1963).
 20. E. T. Copson, *Theory of Functions of a Complex Variable* (Oxford University Press, London, 1935), p. 319.
 21. H. Margenau and G. M. Murphy, *The Mathematics of Physics and Chemistry* (D. Van Nostrand Company, Inc., New York, 1943), p. 292.
 22. MacNair—See discussion following the paper of Ref. 3.
 23. E. K. Wong, "Transit-Time Modulation of Variable Delay Lines," *Technical Report No. 1610-1*, Stanford Electronics Laboratories, Stanford University (May, 1962).
 24. H. G. Slottow, Thesis Research on Plasma Delay Line, Coordinated Science Laboratory, University of Illinois (private communication).
 25. J. Walton, "Stylus Mass and Reproduction Distortion," *J. Audio Eng. Soc.* 11, 104 (1963).

THE AUTHOR



Duane H. Cooper was born in Gibson City, Illinois on August 21, 1923. He received the B.S. degree in 1950 and the Ph.D. degree in 1955, both from the California Institute of Technology, Pasadena.

Since 1954 he has been with the Coordinated Science Laboratory (then the Control Systems Laboratory) of the University of Illinois, Urbana, where he is Research Associate Professor of Physics and Electrical Engineering. In recent years his principal interests have been statistical detection, information and decision theory, with applications to pattern recognition, and the applications of digital computing systems in instrumentation and simulation of physical systems by Monte Carlo methods.

Dr. Cooper is a member of the American Physical Society and of the Audio Engineering Society.

Measurement of Distortions Due to Vertical Tracking Angle Errors in Stereodisk Systems*

J. B. HALTER AND J. G. WOODWARD

RCA Laboratories, Princeton, New Jersey

A laboratory-constructed pickup permitted adjustment of the vertical tracking angle from -25° to $+30^\circ$. Using 400 + 4000 cps intermodulation bands on stereodisks the harmonic distortion, the amplitude modulation, the IM sideband percentages, and the frequency deviation were measured as a function of the vertical tracking angle. While the results of these measurements are in agreement with some of the theoretical relationships which have been developed for tracking-angle phenomena, certain other aspects of the measurements cannot be explained in terms of existing theory and, hence, indicate a gap in our understanding. In particular, the minimum values for the various measured quantities do not all occur for the same tracking angle. This result has practical implications for the measurement of vertical tracking angles by playback techniques and for the design of playback systems for minimum tracking-angle distortion.

INTRODUCTION

THE current interest in the vertical-tracking-angle problem in stereodisk systems has resulted in a series of theoretical and experimental studies^{1,2,3} which have greatly increased our understanding of the phenomena arising from tracking-angle errors. Consideration of intermodulation products, such as are obtained from use of a 400 + 4000 cps test signal, have been especially enlightening. It is known that a tracking-angle error will cause the higher frequency to be both amplitude modulated and phase, or frequency, modulated by the lower frequency of a 2-component intermodulation test signal as played back from a vertically-recorded stereodisk. Consequently, a number of factors are available for experimental measurement and study. In the case of a 400 + 4000 cps test signal these factors include the frequency deviation and the amplitude modulation of the 4000 cps component, the magnitudes of the first-order intermodulation sidebands at 3600 and 4400 cps relative to the 4000 cps carrier, and the harmonic distortion of the 400 cps component. It has been found instructive to measure each of these quantities as a function of the vertical tracking angle. Some measurements of this sort were reported earlier.^{2,3} The present measurements are

much more extensive in that all of the factors listed are included, a greater range of angular variation is used, more data points are used in plotting each curve, and a greater variety of test recordings made under various recording conditions are used.

The precision of the measurements is such that some of the theoretical relationships between the various factors can be checked with fair accuracy. However, the data show some features which have not yet been explained in terms of existing theory. It is hoped that a presentation of these data at this time will stimulate further study and that a suitable and complete explanation of the phenomena will be forthcoming soon.

In order to obtain the measurements required in the present experiments a pickup whose vertical tracking angle could be varied over a wide range, including angles near zero degrees, was needed. Since no commercially-available pickup could meet this requirement, a special laboratory pickup was constructed. The principal characteristics of this pickup and its variable-angle mount will be described briefly in the next section.

VARIABLE-ANGLE PICKUP

The laboratory-constructed stereo pickup used in this work was a variable-reluctance type. The variable-reluctance configuration makes it easier to design for a very small vertical tracking angle and still permits adjustment to large angles by tilting the pickup. The stylus arm was made of a bar of silicon iron with a rectangular cross section, straight throughout its length except for a wider portion where the stylus was mounted between the stationary pole pieces of the pickup. The stylus was mounted

* Presented October 18, 1963 at the Fifteenth Annual Fall Convention of the Audio Engineering Society, New York.

1. Duane H. Cooper, "Tracking Distortion as Phase Modulation," *IEEE Trans. on Audio AU-11*, 41 (March-April, 1963).

2. Benjamin B. Bauer, "The Vertical Tracking Angle Problem in Stereophonic Record Reproduction," *IEEE Trans. on Audio AU-11*, 47 (March-April, 1963).

3. J. G. Woodward and E. C. Fox, "A Study of Tracking-Angle Errors in Stereodisk Recording," *IEEE Trans. on Audio AU-11*, 56 (March-April, 1963).

in a hole near the end of the stylus arm with only enough of the stylus protruding to prevent contact between the end of the arm and the record surface when the stylus was seated in a groove. The other end of the arm was rigidly anchored to the case of the pickup by a length of phosphor-bronze wire. The response-frequency characteristic was flat within ± 2 db between 20 and 20,000 cps, and was determined by the dimensions of the stylus arm and by the dimensions and positioning of rubber damping blocks. The lateral compliance was 11×10^{-6} cm/dyne. The vertical compliance was 5.2×10^{-6} cm/dyne. A vertical tracking force of 4.5 g was found adequate for tests reported here.

The pickup was mounted in a special head which allowed easy variation of the vertical tracking angle. It was



FIG. 1. The laboratory-constructed pickup mounted in a head which permits variation of the vertical tracking angle between -25° and $+30^\circ$.

fastened to a clamp which could be slid along two circular guide rails to permit rotation of the pickup about a horizontal axis perpendicular to the axis of the stylus arm. The axis of rotation passed through the stylus tip and was along a radius of the record when the stylus was riding in a groove. Hence, changes of the vertical angle did not affect the lateral tracking angle or the overhang of the pickup. The circular guides bear scale markings indicating the vertical tracking angle in degrees. The pickup and its adjustable mounting are shown in Fig. 1.

Both positive and negative angles could be used with the pickup and mount just described. For negative angles the pickup and its clamp were slid off the ends of the circular guides, turned through 180° , and replaced on the guides. In this case the record groove moved past the pickup from the front to the rear of the pickup—the reverse of the usual pickup-groove arrangement. This reversal could be made without loosening any mounting screws or removing any leads. The pickup performed satisfactorily in the reverse orientation for negative angles as great as -25° . Hence, a continuously-variable adjustment of angle could

be made from -25° to $+30^\circ$, with only the region between -3.5° and $+3.5^\circ$ being inaccessible due to interference between the pickup case and the record surface.

The pickup could be connected for use in the stereo, the lateral, or the vertical mode of operation. All data to be reported here were obtained in the vertical mode.

MEASUREMENTS

A number of IM test records containing 400 + 4000 cps signals in vertically-modulated grooves were used, some being commercial test records and others being special laboratory recordings. With the exception of one lacquer record, all test records were vinyl pressings. The vertical-mode output of the pickup when playing a record was amplified and fed to the various measuring instruments to obtain data such as are shown in Fig. 2, where the various

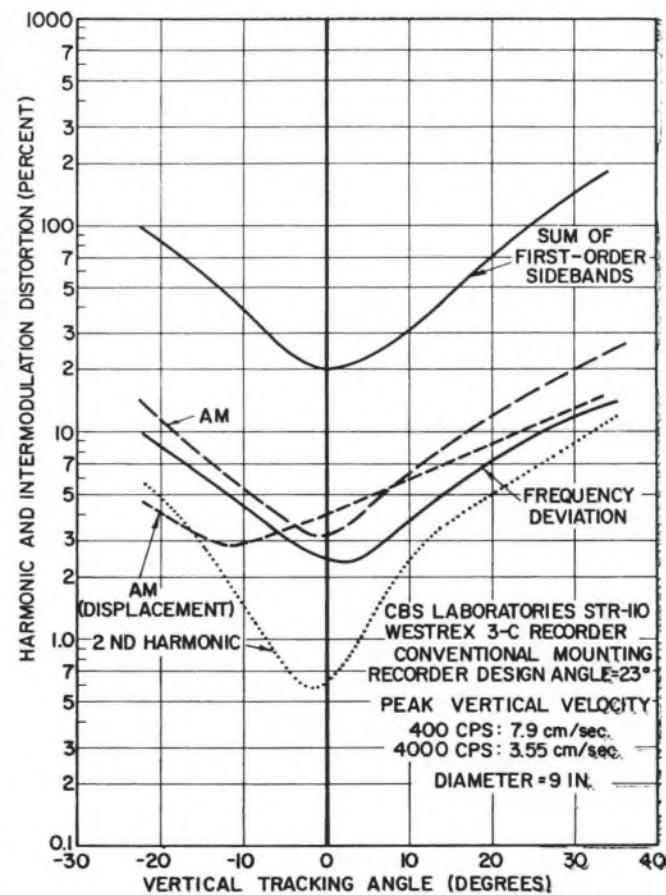


FIG. 2. Measured values of distortion products as a function of the vertical tracking angle in playback. Except as noted, the measurements were made with a velocity-responsive pickup.

IM and harmonic distortions are plotted as a function of the vertical tracking angle. The measuring instruments consisted of a Hewlett-Packard 300A Harmonic Wave Analyzer for measuring % IM sidebands and % 2nd harmonic distortion, an Audio Instruments model 167 Intermodulation Meter for measuring % AM, and a General Radio type 1142-A Frequency Meter and Discriminator together with a Ballantine 305A Peak-Reading Voltmeter

for measuring the % FM peak frequency deviation. At all times the pickup output waveform was monitored on an oscilloscope. The variable-reluctance pickup is a velocity-responsive device, and most measurements reported here are based on this type of response. However, an integrating network could be inserted in the system to give voltages corresponding to a displacement-responsive pickup, and some measurements made on this basis will also be shown.

The results of measurements of 5 records are plotted in Figs. 2 through 6. In each case the abscissa gives the vertical tracking angle on a linear scale and the ordinate gives the various intermodulation and distortion products, measured in percentages, on a logarithmic scale. The percentage first-order sidebands is obtained by adding the magnitudes of the two sidebands, dividing by the magnitude of the 4 kc/sec carrier, and multiplying by 100. The percentage peak frequency deviation is 100 times the maximum swing of the carrier frequency away from its unmodulated value of 4 kc/sec, divided by 4000. The 2nd harmonic distortion and % AM are measured in the customary manner. Unless labelled otherwise, all curves are as obtained with a velocity-responsive pickup.

The data shown in Fig. 2 were obtained from measurements of Group 3B, Band 2, of CBS Laboratories Test Record STR-110. This record was cut with a Westrex 3-C

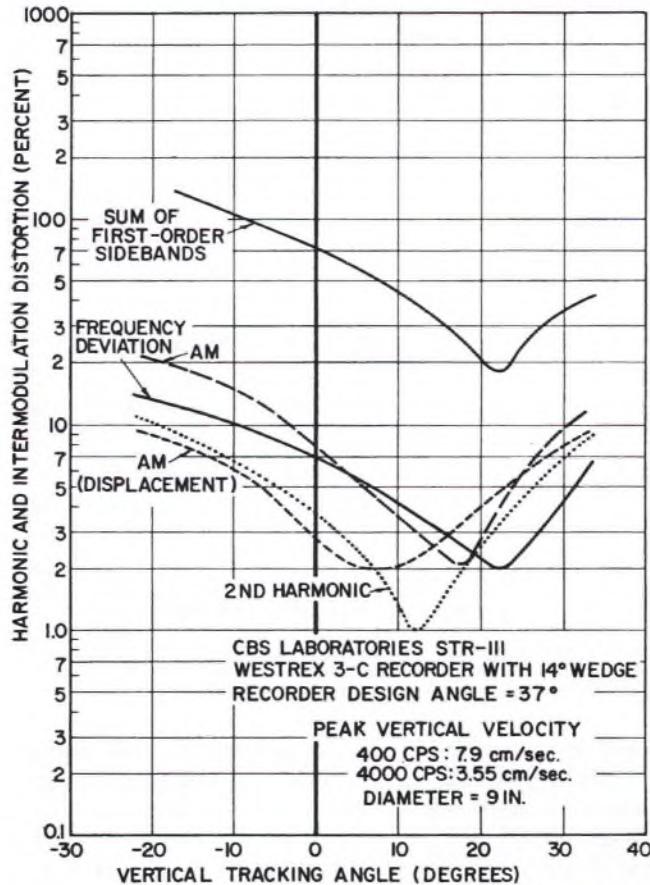


FIG. 3. Measured values of distortion products as a function of the vertical tracking angle in playback. Except as noted, the measurements were made with a velocity-responsive pickup.

Recorder mounted in the normal manner. The reputed vertical angle of the cutter (recorder design angle) in this case is 23°. The effective vertical recorded angle on this record is said to be near 0 degrees.⁴ These results will be discussed in a later section.

Figure 3 shows the results for the same band on CBS Laboratories Test Record STR-111. This record was made under the same conditions as those for STR-110 except that the cutter was tilted forward by 14° to give a vertical angle in recording of 37°. The effective vertical recorded angle on this record is said to be about 15 degrees.⁴

Figure 4 shows the results for an experimental record made at the RCA Laboratories with a Westrex 3-C Re-

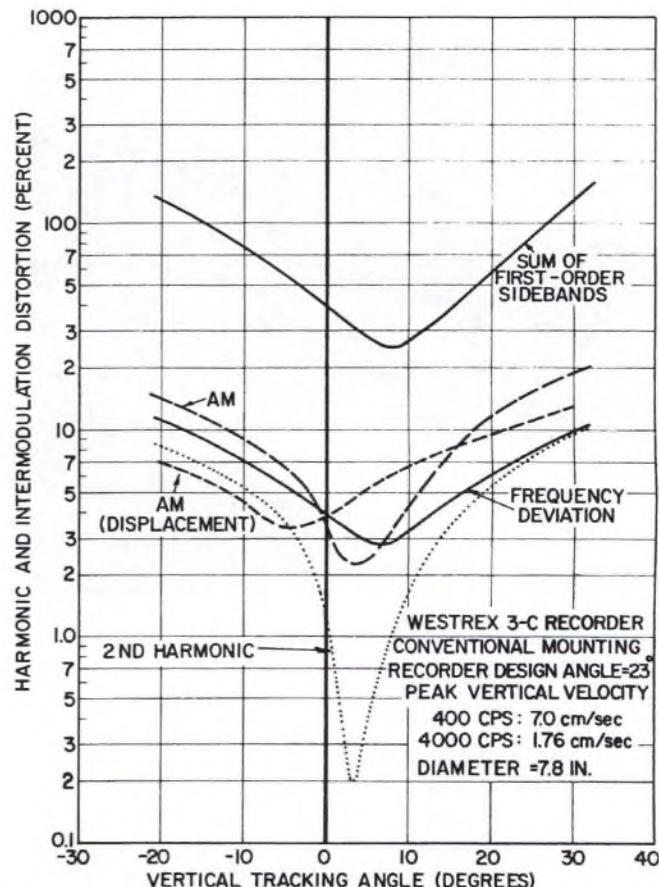


FIG. 4. Measured values of distortion products as a function of the vertical tracking angle in playback. Except as noted, the measurements were made with a velocity-responsive pickup.

corder mounted in the conventional manner with a recorder vertical design angle of 23°.

Figure 5 shows the results for an RCA Laboratories record made with the Westrex 3-C Recorder tilted forward by 10° to give a recorder vertical design angle of 33°.

Finally, Fig. 6 shows the results for a record made for the RCA Laboratories using a Teldec recorder having a recorder vertical design angle of 0°. This was a master lacquer record which was beginning to show signs of wear

4. B. B. Bauer, *op. cit.*

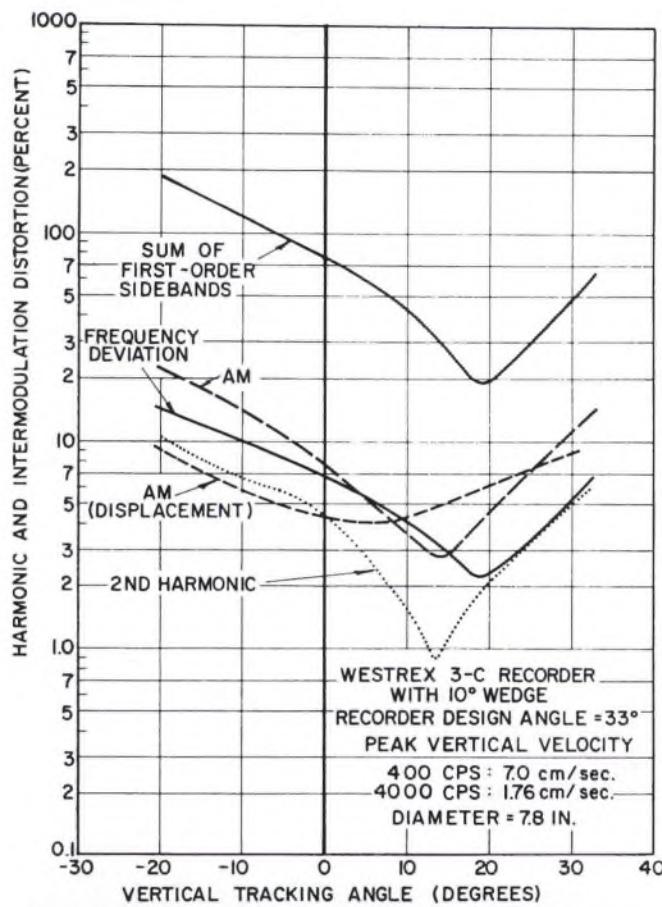


FIG. 5. Measured values of distortion products as a function of the vertical tracking angle in playback. Except as noted, the measurements were made with a velocity-responsive pickup.

when these measurements were made. Consequently, the minimum values observed for some of the products are somewhat greater than they were when the record was new. However, the general shapes of the curves and the angles at which the minima occur have not changed appreciably.

Before proceeding with a discussion of the results presented in Figs. 2 through 6, the effects of record wear will be considered briefly, since wear was found to affect some of the measurements.

EFFECTS OF RECORD WEAR

In general, intermodulation products increase in magnitude as record wear progresses. Care was taken in the present tests to replace a record before wear could influence the results, the one lacquer record mentioned above

TABLE I. Effect of record wear on intermodulation products.

	Initial value	After 22 plays	After 50 plays
		% —	
1st-order sidebands	26	26	27
Frequency deviation	2.8	2.8	2.8
2nd harmonic	0.8	0.4	0.5
AM (velocity)	2.3	3.8	4.2
AM (displacement)	5.5	6.8	7.2

being the only exception since no replacement was available. The effect of record wear was investigated systematically by measuring the IM products on a new record and again after the same band on the record had been played a number of times. The results of one such test are presented in Table I. The record used was that from which the data of Fig. 4 were obtained.

All of the readings in Table I were taken for a vertical tracking angle of $+6^\circ$, which is close to the value corresponding to the minima in the curves of Fig. 4. However, the repeated playings to produce the wear were made with the pickup set at an angle of $+25^\circ$ since it was thought that the wear might be more severe at the larger angles for which the vertical component of compliance of the stylus arm is reduced. Two other series of wear tests made with a 6° angle and a -13° angle showed that the pickup angle did not alter the amount of wear significantly.

The data of Table I indicate that the record wear increased only the AM intermodulation products. The % 2nd harmonic appears to change, but the measured values are very small and lie within a range where the distortion due to vertical tracking angle errors is partially obscured by other distortions in the system and where the accuracy of measurement is reduced. Consequently, the tabulated

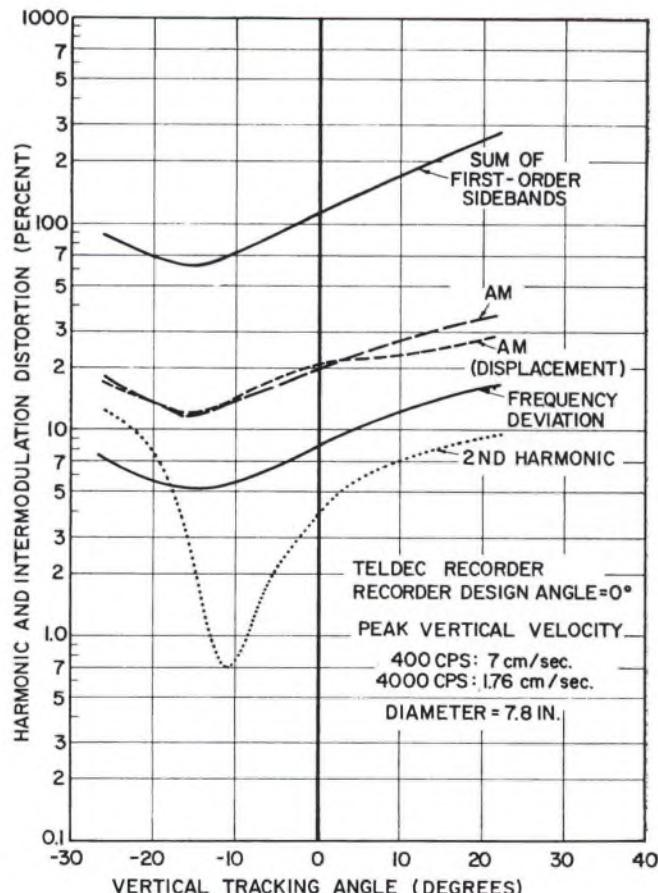


FIG. 6. Measured values of distortion products as a function of the vertical tracking angle in playback. Except as noted, the measurements were made with a velocity-responsive pickup.

change is not judged to be significant. It is reasonable to find that only the AM products are changed, since the principal effect of wear is to reduce the amplitude of the high-frequency component on the positive peaks of the low-frequency component in the groove. The large displacement amplitude of the low-frequency component is not easily changed by playing the record with a well-designed pickup. When measurements such as these are made with a vertical tracking angle considerably removed from that corresponding to minima in the curves, e.g., 25° in the present case, the IM products due to the tracking-angle error are so large as to completely obscure the effects of record wear. Thus, we find that a moderate amount of wear becomes an important consideration only in a playback system which has been adjusted to minimize tracking-angle distortion.

DISCUSSION OF RESULTS

Each of the IM and distortion products in Figs. 2 through 6 is observed to have a definite minimum at some value of the vertical tracking angle. In every case the vertical tracking angle corresponding to a minimum is considerably smaller than the vertical design angle of the recorder used to cut the record. This discrepancy is associated with the recording process and has been discussed previously.⁵ The curves of the percentage of the products as a function of angle are approximately symmetrical about their minima, this being in accord with elementary theory. At the minimum of each curve the contribution of the vertical-tracking-angle error to the measured distortion is very small, and the residual minimum distortion arises from other sources. In the case of the sum of the first-order sidebands and of the frequency deviation the residual distortion is principally tracing distortion. Formulas for calculating the % FM sidebands due to tracing distortion are available,^{6,7} and when such calculations are made for the conditions applicable to the present tests, the calculated values for the sum of the % 1st-order sidebands are in good agreement with the measured minimum values in the various sets of curves. By means of FM theory the calculated values of 1st-order sidebands can be used to calculate the frequency deviation. When this is done, the theoretical values of % deviation due to tracing distortion are also in good agreement with the values actually measured at the minima of the curves. It has been shown⁶ that tracing distortion is a phase-modulation phenomenon and, hence, should exhibit the FM effects just described but should not produce AM on a displacement basis to any pronounced degree. The

minimum value of the displacement AM is, therefore, determined by other factors such as non-linearities in various parts of the record-playback system not associated with the groove-stylus relationships. The measured minimum values of a very few percent of displacement AM are, indeed, within the range which is present in the recording amplifier, the cutter and other elements of the system. This is also true in the case of the 2nd harmonic distortion.

Theoretical relationships between the frequency deviation and the FM and AM sideband amplitudes seem to be generally supported by the measurements. This will be illustrated by one typical example from the data of Fig. 4. For a data point at -19° the % frequency deviation ($\Delta f/4000 \times 100\%$) was measured to be 10.5%, which corresponds to a modulation index of $m_f = \Delta f/400 = 1.05$. Using this value for the argument of the 0th- and 1st-order Bessel functions, tables of these functions give the magnitude of each 1st-order FM sideband as 61% of the 4 kc/sec carrier, or 122% for the sum of the two sidebands. The measured value was 124%. In measurements of this type it has been consistently observed that the upper sideband at 4.4 kc/sec is greater in magnitude than the lower sideband at 3.6 kc/sec. In the present example the upper sideband was 72% and the lower sideband was 52% of the carrier amplitude as measured with a velocity-responsive pickup. This difference in magnitude is also in accord with theory, as will now be shown.

The difference in magnitude of the upper and lower side-

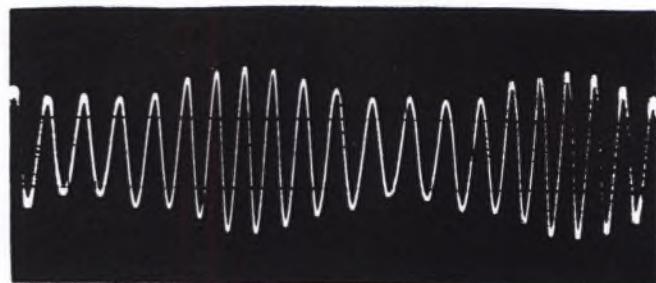


FIG. 7. Oscilloscope trace of the 4 kc/sec component in playback of a 400 + 4000 cps, vertical-cut signal. Both AM and FM are generated as a result of a vertical-tracking-angle error. Note that the maximum frequency and the maximum amplitude occur at the same time.

bands is due partly to the use of a velocity-responsive pickup in the measurement of the sidebands and partly to AM, which is known to exist on a displacement basis. Since a velocity-responsive pickup was used, the output of the pickup is proportional to frequency. If the output of the velocity-responsive pickup were passed through an integrating circuit or if a displacement-responsive pickup had been used, the measured sideband amplitudes would have been $72 \times 4000/4400 = 65\%$ for the upper sideband and $52 \times 4000/3600 = 58\%$ for the lower sideband relative to the carrier. The remaining discrepancy between these values and the value of 61% calculated above from the frequency deviation is due to displacement AM.

Both theory and observation of the 4 kc/sec waveform show that the maximum frequency due to FM and the

5. B. B. Bauer, *op. cit.*; J. G. Woodward and E. C. Fox, *op. cit.*

6. E. C. Fox and J. G. Woodward, "Tracing Distortion—Its Cause and Correction in Stereodisk Recording Systems," *J. Audio Eng. Soc.* 11, 294 (1963).

7. M. S. Corrington and T. Murakami, "Tracing Distortion in Stereophonic Disc Recording," *RCA Review* 19, 216 (Jan., 1958). In this reference the applicable formulas are shown as yielding % IM. However, as the text makes clear, it is actually the sum of the 1st-order sidebands as a percentage of the carrier which is calculated. There is also some question regarding the proper value of the low-frequency velocity to be used when these formulas are applied to the case of a vertically-modulated groove.

maximum amplitude due to AM occur at the same time. This can be seen in Fig. 7 which shows an example of a frequency- and amplitude-modulated carrier resulting from playback of a 400 + 4000 cps signal with a vertical tracking angle error. The amplitude modulation observed is due to both the displacement AM and the AM caused by sensing a frequency-modulated signal with a velocity-responsive pickup. Theory shows that these two amplitude modulations are equal in magnitude and in phase. This phase relationship between the FM and the AM means that the

$58\% = 3\%$. The % AM is equal to the sum of the two AM % sidebands and is 7%. The measured value as taken from Fig. 4 is 6.6%. In view of the inaccuracies in the various measurements, and remembering that higher-order FM sidebands were not taken into account, the agreement found here is better than one has a right to expect. It should be understood that the measurement will not confirm the vertical-tracking-angle theory in the vicinity of minima in the various curves since other sources of distortion are effective here. Also, other disturbing factors still to be discussed are more serious close to the minima.

Such qualitative and quantitative agreement between theory and measurement encourages us to believe that most of the existing theory for vertical-tracking-angle phenomena is basically sound. However, let us now consider one of the more striking features of the data presented in Figs. 2 through 6, namely, the different vertical tracking angles for which minima are obtained for the various IM and distortion products in each set of data. All of the theoretical relations of which the present writers have knowledge indicate that all of the IM and distortion products should have their minimum values for the same value of vertical tracking angle. The type of discrepancy observed here was already noted in earlier published results⁸ for which the % sidebands and the % 2nd harmonic had minima at different angles. At that time it was conjectured that the discrepancy might be due to an interaction between the IM and the distortion products due to tracing distortion and those due to a vertical-tracking-angle error. It is now known that tracing effects are not sufficient to account for the observed discrepancy in angle.

The degree to which tracing effects might influence the vertical tracking angle for minimum IM and distortion products was ascertained by recording a test record identical with that used for Fig. 5, except that the Dynamic Recording Correlator⁹ was used to compensate for tracing distortion. When this record was used for a set of playback tests, the angles for the various minima were almost exactly the same as those measured in Fig. 5. Thus, while the presence or absence of tracing distortion may change the minimum values of the various products, it does not appear to affect significantly the angle at which the minimum is observed.

It should be noted that a certain pattern appears to exist in the relative angular locations of the various minima in a given set of data. In most cases the frequency deviation and the 1st-order sidebands have minimum values at the same angle, this angle being greater than those for minima of the other products. The frequency deviation and the sum of the 1st-order sidebands are manifestations of frequency modulation alone. The IM product having its minimum at the smallest angle is the AM for a displacement-responsive pickup. In this case only amplitude modulation on a displacement basis is involved. Theoretically, the velocity-responsive AM is due to displacement AM sidebands and FM-AM sidebands in equal amounts, so it is not surprising

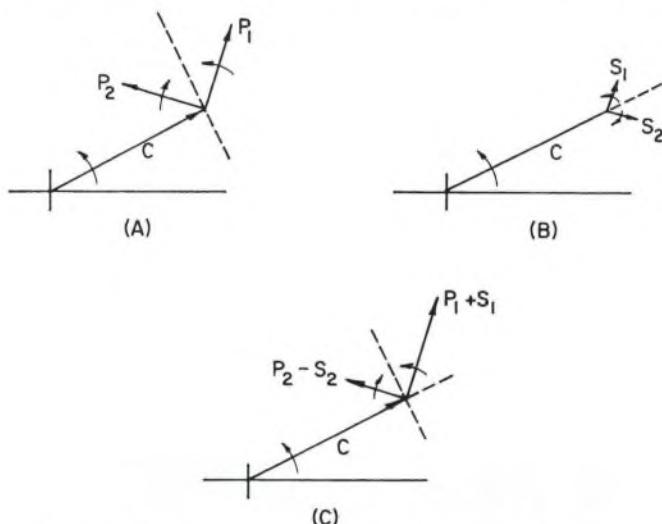


FIG. 8. Vector diagrams representing the relationships between carrier and sidebands in amplitude modulation and phase or frequency modulation. (A) FM case; (B) AM case; (C) Resultant of FM and AM.

upper AM sideband adds to the upper FM sideband, whereas the lower AM sideband subtracts from the lower FM sideband. This latter fact may be deduced from the vector diagrams frequently used to depict amplitude and frequency or phase modulation, as shown in Fig. 8. In these diagrams the unmodulated carrier is represented by a rotating vector, C, which is situated as shown at a particular instant. In Fig. 8A the two counter-rotating vectors, P_1 and P_2 , representing upper and lower sidebands, are shown with the phase relationship which results in frequency or phase modulation. In Fig. 8B the counter-rotating vectors, S_1 and S_2 , are shown with the phase relationship which results in amplitude modulation. The phase relationship between the AM and the FM sideband vectors as drawn in Figs. 8A and 8B is the one which meets the requirement that the frequency and amplitude have their maximum values at the same instant. For this condition it is seen that the resultant lower sideband is $P_2 - S_2$ and the resultant upper sideband is $P_1 + S_1$, as shown in Fig. 8C. We note that the sum of the two resultant sidebands is equal to the sum of the FM sidebands only, and is independent of AM effects.

With this knowledge we can return to the numerical example and calculate the % AM which we would expect to measure with a displacement-responsive pickup. The discrepancy for the upper sideband found earlier is $65\% - 61\% = 4\%$. The discrepancy for the lower sideband is $61\% -$

8. J. G. Woodward and E. C. Fox, *op. cit.*

9. E. C. Fox and J. G. Woodward, *op. cit.*

to find that the minimum for the AM of a velocity-responsive pickup is intermediate between the pure displacement AM and the pure FM cases. The angle for the minimum of the 2nd harmonic is usually close to, but slightly smaller than, the angle for the minimum of the velocity-responsive AM. This pattern may furnish a clue to the elaboration of the theory required to explain the observed results.

PRACTICAL IMPLICATIONS AND CONCLUSIONS

Since the various intermodulation and distortion products are found to have their minimum values at different angles, the question immediately arises as to which of the angles, if any, is a true measure of the effective vertical recorded angle on the record. This question cannot be answered conclusively until our understanding of the phenomena is more complete. However, it is tempting to believe that the frequency-modulation effects (1st-order sidebands and frequency deviation) come closest to providing a correct measure of the recorded angle. This is because the FM effects are dependent on only the slope of the large-amplitude, relatively stable low-frequency component of the two-component IM test signal, whereas the AM effects are somewhat more sensitive to the amplitude and to amplitude changes of the high-frequency component. Both the frequency deviation and the sum of the 1st-order sidebands are unaffected by AM, and good agreement is consistently obtained between the two completely-independent techniques by which they are measured.

The question regarding the true value of the vertical recorded angle has practical importance because a particular 400 + 4000 cps IM test record is being seriously considered for use in a standard procedure for measuring the vertical tracking angles of pickups. This procedure, as tentatively proposed, would be based on the measurement of the % amplitude modulation of the 4 kc/sec component, and would not specify whether the measurement is to correspond to a velocity-responsive or to a displacement-responsive pickup. It appears that some additional study is desirable before a standard procedure is finally adopted. A conventional IM meter can be used conveniently to measure the % AM. A frequency meter and discriminator is no less convenient to use for measuring the % frequency deviation, and this may turn out to yield a more correct measurement for determining the vertical tracking angle of a pickup by means of playback tests.

Another practical question arises when one wishes to optimize a playback system to produce the minimum amount of audible distortion. The relative subjective seriousness of AM effects and FM effects in music reproduced from stereodisks cannot now be stated with certainty. However, a reasonable compromise would seem to call for a vertical tracking angle intermediate between the angles for which minima occur in the displacement-responsive AM and in the frequency-deviation characteristics. It is necessary, of course, to know that the test records used for the IM measurements were made under the same conditions as those under which the commercial records containing musical recordings were made. The measurements plotted in Fig. 6

for the Teldec recorder represent the situation for a number of releases by European record companies and by some American companies. A noticeable reduction in audible distortion in reproduced music from some of these records was achieved by using a pickup having a vertical tracking angle of -15° . This angle was obtained by mounting the pickup backwards on the pickup arm. The data shown in Figs. 2 and 4 are representative of the conditions applicable to the great majority of American releases and some European releases prior to early 1963. For these records a pickup having a very low positive vertical tracking angle (say about 5°) would provide the lowest distortion. The conditions represented in Fig. 3 are said to apply to Columbia Records' stereodisk releases since about January 1963. The conditions represented by Fig. 5 apply to the DYNAGROOVE stereodisk releases introduced by RCA Victor in March 1963. In both these latter cases an effort was made to adhere to a proposed standard vertical recorded angle of 15° , and the results achieved represent a significant step toward closer control of the recording conditions. It is hoped that standardization of the vertical angle in both recording and playback can become industry-wide in the near future.

THE AUTHORS



J. G. WOODWARD



JEROME B. HALTER

J. G. Woodward received a B.S. degree in 1936 from North Central College, a M.S. in physics in 1938 from Michigan State University, and a Ph.D. in physics in 1942 from The Ohio State University. In March, 1942, he joined the engineering activity of the RCA Victor Division in Camden, New Jersey and later that year moved to the newly formed RCA Laboratories in Princeton, New Jersey, where he has since been engaged in research in audio and electro-acoustics.

Dr. Woodward is a senior member of the Institute of Electrical and Electronics Engineers, a member of the IEEE Professional Group on Audio, a member of Sigma XI, and a Fellow of the Acoustical Society of America, the American Association for the Advancement of Science, and the Audio Engineering Society.

Jerome B. Halter received a B.S. degree in 1949 and a M.S. degree in 1950 from Washington University, St. Louis, Missouri. In September, 1950, he joined the Technical Staff of RCA Laboratories in Princeton, New Jersey where he has been engaged in research in tape and disk recording, and sound reproduction.

Mr. Halter is a member of the Audio Engineering Society, the Professional Group on Audio of the Institute of Electrical and Electronics Engineers, Sigma Xi, Tau Beta Pi, and the Acoustical Society of America.

ON CUTTING-ANGLE COMPENSATION

DUANE H. COOPER

University of Illinois, Urbana, Illinois

At the fifteenth Annual Fall Convention of the Audio Engineering Society, Mr. J. B. Halter presented new measurements made at RCA of vertical tracking angle distortion.¹ The data in this paper [on p. 8 of this issue of the *Journal*] seem to show an 8° change in effective cutting angle as the cutting diameter was changed from 9 in. to 7.8 in. Such a dependence upon recorded wavelength, though already known, had not been thought to be so extreme up to now.

Also, the data showed amplitude intermodulation effects which could not be entirely explained by linear phase automodulation theory.² For example, the minimum in amplitude intermodulation did not come at the same angle as the minimum in phase intermodulation. A similar discrepancy is seen with the angle as measured by harmonic content.³

These discrepancies could account, in part, for the choice at RCA of a 10° mounting wedge, for use with the Westrex 3-C cutter, and the choice at CBS of a 14° wedge.⁴ The data indicate that either choice is a compromise among a broad range of possible angles, so that one would be hard put to decide which compromise is the more nicely drawn.

More than that, these valuable data suggest that there are elastic processes in the cutting action that have not yet been fully explored. One would like to see whether the moving parts of the Westrex cutter might be stiffened, and whether such modifications would result in more consistent data. Alternatively, other cutters, such as the Teldec or Ortofon, should be examined to see if they might show smaller elastic effects than in the Westrex. If so, it would be interesting to see if the use of such cutters would lead to more consistent data. In this way, elastic processes in the cutter and elastic processes in the lacquer might be isolated.

It should be emphasized that if this variable angle performance should be shown to be an inherent feature of the cutting process, then the proper maintenance of the effective cutting angle is likely to require the use of complicated procedures. There is not only the broad spectrum of wavelengths inherent in the audio band to be considered in maintaining a standard effective cutting angle, but also the further variation of wavelength with cutting diameter. Electronic adjustment of effective cutting angle is a theoretical possibility, but the adjustment of such an apparatus to take account of the wavelength effect, based on further empirical data, would still make for a substantial complication.

It has been shown that the MacNair-Hunt re-recording scheme is capable not only of compensation for tracing error, but also of providing an automatic adjustment of effective cutting angle to match the tracking angle of the cartridge used in playback for re-recording.⁵ The data showing the removal of tracing error by this means, as given by Fox and Woodward, are very promising, especially since no evidence was presented to suggest that great care was exercised to normalize levels and to minimize tip-mass distortion in the playback.⁶ By this method, any effective cutting angle is brought into agreement with the tracking angle of the playback cartridge automatically, and tracing compensation is thrown into the bargain. Thus, one would expect the variations in angle with wavelength to be removed as well.

Theoretical analysis shows a second-order effect arising from an interaction between the mechanisms causing tracing error and tracking error.⁷ Thus, if tracing correction is imposed synthetically by such a device as the RCA Dynamic Stylus Correlator, and the resulting signal used to drive an inclined cutter, there is a residual distortion, because these two compensation schemes are not invoked in the proper sequence. Such would be the case even for ideal cutting action different from vertical. The MacNair-Hunt scheme also should show a residual distortion, having the same theoretical source, and of the same order of magnitude.

The same analysis shows that the residual distortion for the

MacNair-Hunt scheme should vanish for a cutter producing a negative effective cutting angle, but of the same magnitude as the tracking angle in the playback cartridge.⁷ In this configuration, also, there are no renormalization problems; it is only necessary that the original cutting levels be accurately maintained in the second cut. Thus, if a 15° cartridge were used in the playback for re-recording and if the natural effective cutting angle of the Teldec or Ortofon cutter, say, could be shown or made to stay, near 15° negative, then the least residual distortion of all should result with those cutters.

An objection to the MacNair-Hunt compensation scheme is that it is a two-step process. The objection may be partially overcome by pointing out that playback and re-recording may proceed simultaneously with the first cutting. All that is required is that the second cutting set-up should have characteristics as nearly identical with the first as possible.

It would be interesting to see data like that presented by Mr. Halter, but obtained for the MacNair-Hunt scheme using a cutter showing the requisite negative angle.

REFERENCES

1. See J. B. Halter and J. G. Woodward, "Measurement of Distortions Due to Vertical Tracking Angle Errors in Stereodisk Systems," *J. Audio Eng. Soc.* 12, 8 (Jan., 1964).
2. Duane H. Cooper, "Unified Analysis of Tracing and Tracking Error," *IEEE Trans. on Audio AU-11* (July-August, 1963).
3. J. G. Woodward and E. C. Fox, "A Study of Tracking Angle Errors in Stereodisk Recording," *IEEE Trans. on Audio AU-11*, 57 (1963).
4. B. B. Bauer, "The Vertical Tracking Angle Problem in Stereophonic Record Reproduction," *IEEE Trans. on Audio AU-11*, 47 (1963).
5. Duane H. Cooper, "Compensating for Tracing and Tracking Error," *J. Audio Eng. Soc.* 11, 318 (1963).
6. E. C. Fox, and J. G. Woodward, "Tracing Distortion—Its Causes and Correction in Stereodisk Recording Systems," *J. Audio Eng. Soc.* 11, 294 (1963).
7. Duane H. Cooper, "Integrated Treatment of Tracing and Tracking Error," *J. Audio Eng. Soc.* 12, 2 (Jan., 1964).

LETTERS TO THE EDITOR

Note to Members: This is your column. It is designed for the discussion of papers published in the Journal and other pertinent topics about which you feel strongly.

FURTHER COMMENTS ON THE MEASUREMENT OF DISTORTIONS IN STEREOGRAPHIC SYSTEMS

J. B. HALTER

RCA Laboratories, Princeton, New Jersey

These comments are in answer to Professor Duane H. Cooper's Letter to the Editor entitled "On Cutting Angle Compensation" (*J. Audio Eng. Soc.* 12, 67, Jan. 1964) in which he mentions a possible dependence of the vertical recorded angle on wavelength. He refers to recent measurements of vertical tracking angle distortion made at RCA Laboratories which were presented at the 15th Annual Convention of the A.E.S. and which also appeared in the January issue of the *Journal*.¹

It was noted in this paper that of the various types of distortion measurements investigated, the frequency deviation and the sum-of-the-first-order-sideband measurements are the least affected by other system distortions and record wear. Thus, the null point of each sum-of-the-first-order-sideband curve and/or the corresponding frequency deviation curve gives perhaps the best indication of the effective "vertical recorded angle" as obtained by means of playback measurements.

The particular measurements referred to by Professor Cooper are those of the CBS Laboratories STR 110 test record which were measured at a 9 in. diameter and of the RCA Laboratories test record which were measured at a 7.8 in. diameter. Both of these records were cut by means of a Westrex 3-C recorder without the use of a wedge for modifying the 23° recorder design angle. The frequency deviation and the sum-of-the-first-order-sideband measurements of these two records indicate an effective "vertical recorded angle" of about 0° for the CBS record and about 8° for the RCA record. It was not known, when these measurements were made, what portion (if any) of this 8° difference in angles could be attributed to the 15% change in diameter and wavelength. Subsequent measurements have shown that if all other factors remain the same and the groove diameter (and, thus, the wavelength) alone is varied, the change in vertical recorded angle is relatively small.

Measurements taken at 11, 7.8 and 5 in. diameters on another test record are shown in Figs. 1, 2 and 3, respectively. The measurements shown in Fig. 2 are the same as those shown in Fig. 5 of the paper previously referred to.¹ It can be seen from the sum-of-the-first-order-sideband curves or the frequency deviation curves that the difference in angles between the inner and outer groove diameters is only about 2° or 3°.

Thus, the 8° difference which occurs for only a 15% change in diameters must be due, for the most part, to factors other than the groove speed or wavelength. For example, in recording the two records in question, there could have been small differences in the stiffness and shape of the styli, in the seating of the tapered stylus shanks into the stylus holders, in the mechanical linkages of the cutters, in the mechanical properties of the lacquer masters, in the temperature of the styli, etc. One factor which may be of particular importance in causing some variation of the effective "vertical recorded

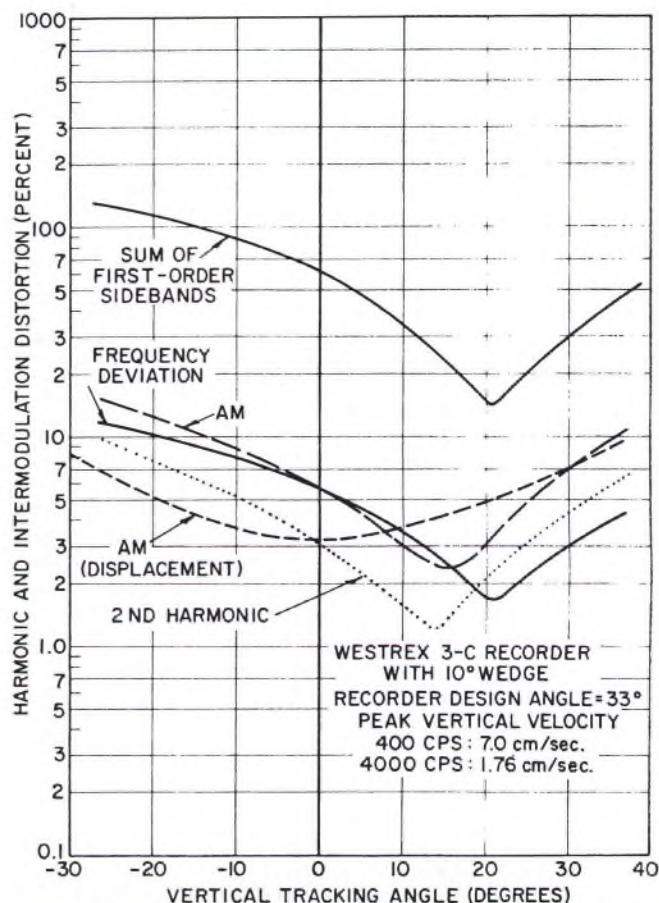


FIG. 1. Measured values of distortion products as a function of vertical tracking angle in playback at an 11 in. diameter. Except as noted, the measurements were made with a velocity-responsive pickup.

angle" is the depth of cut; however, in this particular case the average vertical displacements of both records were nearly the same.

There are several other features shown in Figs. 1, 2 and 3 which may be of interest. It can be seen that there is only a small increase (about 33%) in the broad minima of the AM (displacement) curves from the outer to the inner groove diameters. This is because other system amplitude IM distortions tend to mask small amplitude IM distortions resulting from vertical tracking errors. It can also be seen that the second-harmonic minima actually decrease in magnitude as the diameter decreases due to cancellation effects with other system second-harmonic distortion.

It is the sum-of-the-first-order-sideband curves and the frequency deviation curves which agree closely with theory. Both of these curves are an indication not of amplitude modulation but of fre-

¹ J. B. Halter and J. G. Woodward, "Measurement of Distortions Due to Vertical Tracking Angle Errors in Stereodisk Systems," *J. Audio Eng. Soc.* 12, 8 (Jan., 1964).

quency deviation of the 4 kc test signal produced as a result of tracking-angle errors and, also, tracing distortion. The magnitude of each of the curves at the null point is a measure of the tracing distortion present at that diameter in terms of either the % frequency deviation or the % sum-of-the-first-order-sidebands. It can be seen that the minima of these curves increase by about 4 times and become

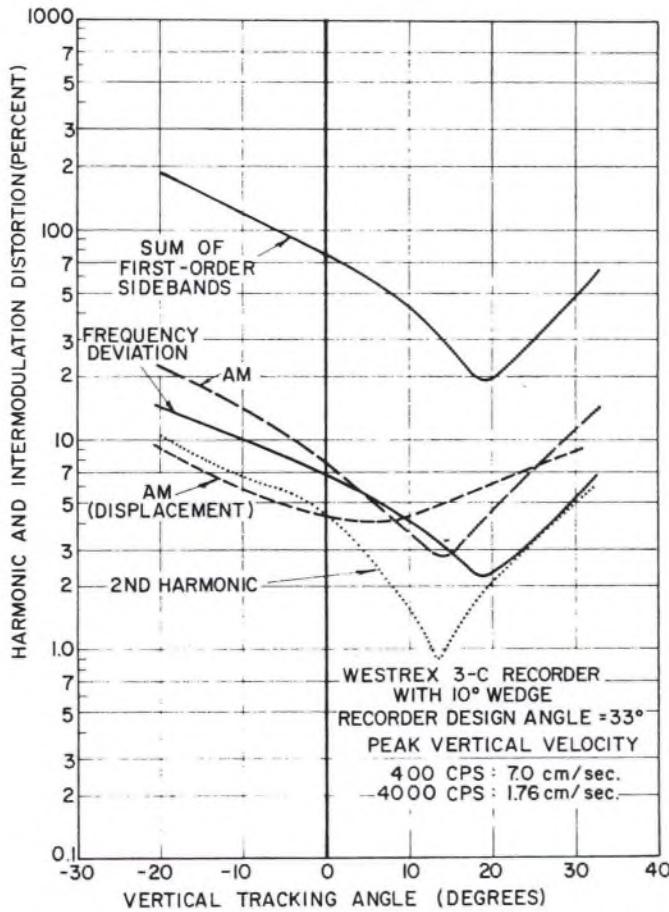


FIG. 2. Measured values of distortion products as a function of vertical tracking angle in playback at a 7.8 in. diameter. Except as noted, the measurements were made with a velocity-responsive pickup.

broader as the groove diameter is approximately halved. The reason for this is that, theoretically, tracing distortion should vary inversely with the square of groove diameter while tracking-angle-error distortion should only vary inversely with the groove diameter. It is clear, nevertheless, that in the case of pickups with vertical-tracking angles larger than about 25°, this type of IM distortion that is due to vertical-tracking errors will exceed that due to tracing distortion even at the inner diameters.

It should be borne in mind, however, that this type of IM test signal (SMPTE) is primarily a test of intermodulation which is produced by signals in the frequency region near 400 cps where the displacements are relatively high. Measurements made with two more closely spaced higher-frequency signals (the CCIF method) show much broader nulls as the vertical tracking angle is varied. This indicates the relatively larger significance of tracing distortion in the higher frequency region.

The reduction of tracing distortion by means of the RCA Dynamic

Recording Correlator has been described in a published paper.² In Fig. 11 of that paper, reproduced here as Fig. 4, the reduction of tracing distortion when using an SMPTE type test signal is shown. These data were obtained using the same test record as was used to obtain the vertical tracking angle data of Figs. 1, 2 and 3 shown here. Thus, the minima of the three sum-of-the-first-order-sideband curves shown here correspond to the first-order sideband data shown in Fig. 4, which were obtained without the use of the Dynamic Recording Correlator. Those curves, however, are given in terms of the separate channel measurements of the two individual sidebands. These left and right channel measurements may be averaged to find the percentages corresponding to a vertical output and then these upper and lower sideband averages added to find the sum-of-the-first-order-sidebands. If this is done, for example, at the 5 in. diameter, one obtains from Fig. 4 a value of about 60%. The minimum value of the sum-of-the-first-order-sideband curve for a 5 in. diameter,

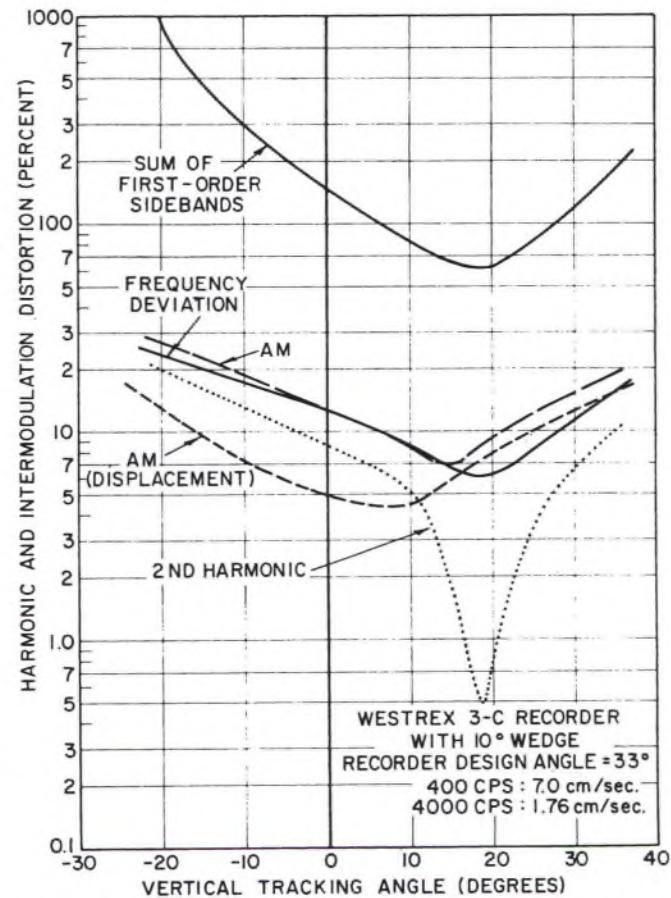


FIG. 3. Measured values of distortion products as a function of vertical tracking angle in playback at a 5 in. diameter. Except as noted, the measurements were made with a velocity-responsive pickup.

shown here in Fig. 3, is about 61% which is in reasonably good agreement. It is convenient to also express the distortion present when the correlator is used in terms of the sum-of-the-first-order-sidebands. This may be done in a similar manner by using the data of Fig. 4 which were obtained with the correlator in use in the system. At a 5 in. diameter, a value of about 17% is found. This would correspond to a frequency deviation of 1.7%.

Thus, in the case of vertical modulation at the inner diameters it is found that the Dynamic Recording Correlator will reduce the

² E. C. Fox and J. G. Woodward, "Tracing Distortion—Its Cause and Correction in Stereodisk Recording Systems," *J. Audio Eng. Soc.* 11, 300 (1963).

tracing distortion about $60/17 = 3.5$ times. Reductions in tracing distortion by a factor of 6 have been obtained with this device under slightly different playback conditions.

At the present time it is difficult to say how much frequency deviation and amplitude IM distortions must be limited in order that they not be perceptible or objectionable to critical listeners. The sidebands produced because of a vertical tracking error are, theoretically, 90° out of phase with regard to those at the same frequencies that result from tracing distortion. Therefore, the distortion produced as a result of a vertical tracking error cannot be used to cancel tracing distortion and vice versa. In order to derive the maximum benefits from reducing either of these two sources of distortion, the other source of distortion must be minimized as well. If not, a large reduction in the distortion produced by one source may not cause a very noticeable change in the total distortion. In addition, it is likely that a reduction in distortion resulting from both of these sources may not be appreciated in the presence of other significant distortion such as that which may be produced if overloading occurs in any portion of the record-reproduce system.

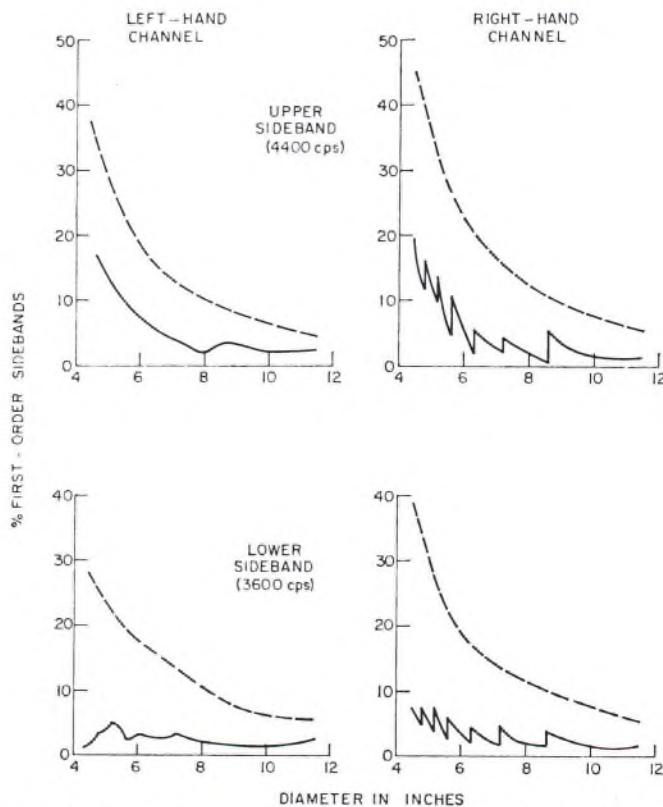


FIG. 4. Playback measurements of first-order sidebands of a 400 + 4000 cps signal recorded without (broken curves) and with (solid curves) the Dynamic Recording Correlator in the system.

On Tracking and Tracing Error Measurements*

DUANE H. COOPER†

University of Illinois, Urbana, Illinois

Anomalies in published distortion data for tracking and tracing error are explained as caused by contamination from other distortions. The two-channel properties of these two errors are studied to reveal ways in which they may be sharply delineated from one another and from spurious effects. A standard tracking error test is devised offering a ten-fold improvement in precision over earlier tests. Tentative results from an experimental recording indicate partial fulfillment of this offering. Several interesting tests for tracing error are proposed.

INTRODUCTION

FIGURE 1 illustrates the difficulties facing one seeking to interpret distortion measurements as indices of track-

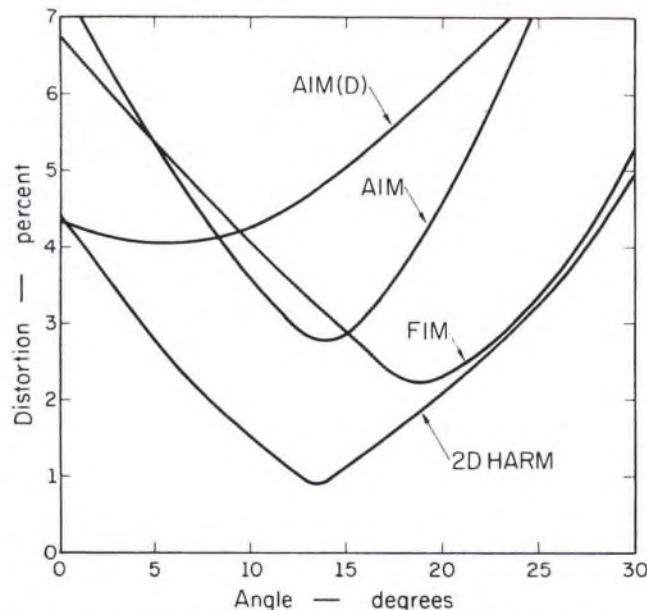


FIG. 1. Vertical tracking error distortion. Four different measures of distortion are shown replotted from the data of Halter and Woodward (Ref. 1). They show second harmonic distortion, frequency intermodulation distortion, and amplitude intermodulation distortion plotted *versus* vertical tracking angle. The measurements were all made in the velocity waveform save the one marked AIM(D) for which a displacement waveform was used. The conditions were: 3.9 in. groove radius, 7.0 cm/sec recording velocity in the vertical channel at 400 cps and 1.76 cm/sec at 4000 cps, Westrex 3-C cutter with 10° mounting wedge.

* This work was done, in part, as consultant to Shure Brothers, Inc., Evanston, Illinois.

† The author's biography appears in the January 1964 issue of the *Journal* on p. 7.

ing angle error in phonograph reproduction. It shows, replotted to linear scales, some of the valuable distortion data published by Halter and Woodward.¹ Four modes of distortion measurement are shown. The values are plotted in percentage units *vs* the vertical tracking angle in degrees, as indicated by the tilt of a playback pickup of known inclination. Presumably, the minimum in a given curve is indicative of a minimum in vertical angle used in cutting the record. These minima are seen to come, for the various curves, at the angles of 6°, 13.5°, 14°, and 19°.

These data were all derived from pressings made from a single master lacquer. Only one cutter was used—the Westrex 3C fitted with a 10° mounting wedge—and only one pickup was used. Halter and Woodward believe that only the 19° value may be trusted. The present investigation will support that belief.

The minimum values of the distortion presumably measures the distortion remaining after tracking error is removed. This presumption is, of course, invalid for those minima which fall at angles not representative of zero tracking error. Nevertheless, the minima are in fair agreement with the theoretical values for tracing distortion for all curves save one, the measurement of amplitude intermodulation on a displacement basis. There should be no measurable tracing distortion for that mode.

If there were only tracing and tracking distortion present, the theoretical shape for these curves would be hyperbolic. The plotted shapes are largely anomalous. More important, the plotted data are inadequate as indices of the nonlinear processes present. The original data also included measurements of the intermodulation product tones obtained by wavemeter analysis. Those measurements indicated distortion levels tenfold greater than shown here. That 10:1 ratio is in agreement with theory.

Figure 2 shows a plot from the same data showing how the angles for minimum distortion vary with cutting radius.

The indicated angles for the various modes disagree with one another, not only at particular radii, but also in the way that they vary, some showing sharp variations. The curves shown are intended to help identify the various modes, not to represent adequate interpolation.

These data emphasize a need for a measurement procedure that is 1. apt, 2. precise, and 3. consistent. Clearly not all of the modes of Fig. 1 are equally apt. The exceedingly large amounts of distortion indicated by multiplying the values of Fig. 1 by the factor 10 emphasize the need for precision. The variation with radius shows that some modes of measurement are not consistent. The establishment of an industry standard for vertical angle demands that these three qualities be provided.

The present paper offers an explanation of the variability of Figs. 1 and 2 as being caused by contamination from

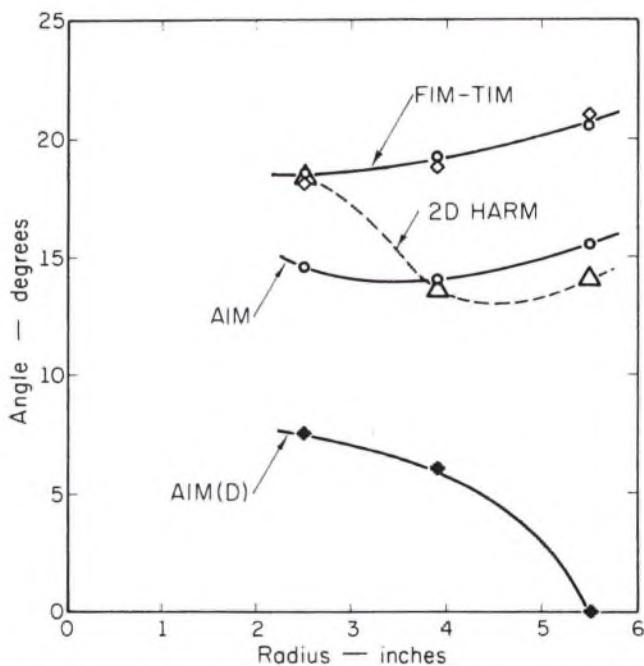


FIG. 2. Vertical tracking angles for minimum distortion as a function of groove radius. The data, including that for total intermodulation (TIM) are taken from Halter (Ref. 1). Except for the TIM data, plotted by open circles and referring to the sum of the first order intermodulation sidetones, the same measures of distortion are used as in Fig. 1. The curves are drawn to identify families of data points. The same recording conditions, except for groove radius, obtain as for Fig. 1.

other sources of distortion. A review of the nature of the various kinds of distortion is offered, showing how they may be classified according to clearly distinguishable properties. The classification first distinguishes amplitude-perturbing nonlinearities from delay-perturbing ones. Then, it invokes the two-channel properties of the stereo record groove to show how the delay-perturbing nonlinearities may be sorted into two classes.

This investigation leads to the devising of a proposed standard test for tracking error which is singularly free of contamination by other errors. The delineation is so sharp as to offer an enhancement of accuracy in angle measure-

ment by an order of magnitude. To the extent that consistency may be assured by aptness and precision, the test should challenge the consistency with which calibration methods may be devised. An experimental recording embodying such a test produces results in partial agreement with the predicted enhancement in accuracy. Full agreement is limited by a residual distortion whose identification must await further experimental work.

The same considerations lead to devising tests for tracing error. Though a substantial improvement over present tests is offered, the extreme purity of the tracking test does not seem to be available. The most promising test, for example, requires lateral tracking error, though not the vertical, to be brought to a null before full accuracy may be obtained. Even then, some contamination from a part of the groove-wall deformation effects may remain. For this reason, and because the need is less pressing, the tests are offered as research tools rather than proposed as standards.

NONLINEAR DISTORTION

Memoryless nonlinear signal transmissive devices may be regarded as ones for which the output signal $f_o(t)$ is a nonlinear function of the input signal $f_i(t)$, expressed as

$$f_o = G(f_i), \quad (1)$$

for each instant of time. The form of G could be one of a great variety, but for a "gentle" nonlinearity, it may be represented as a power series

$$f_o(t) = f_i(t) + af_i^2(t) + \dots \quad (2)$$

as a preliminary step towards further analysis. Equation (2) may also be written as representing automodulation

$$f_o(t) = f_i(t) [1 + af_i(t) + \dots], \quad (3)$$

in which the departure from linearity may be measured by an automodulation index, the product of a by some measure of the magnitude of f_i . If a should happen to vanish, interest would transfer to the coefficient of the next higher non-vanishing term.

When the transmissive device involves memory, the transmission function G will involve derivatives or integrals of the input signal as well. For example, a transmission system producing only a delay is linear but not memoryless. With the help of Taylor's theorem, the transmission function is seen to be linear in f_i and its derivatives,

$$\begin{aligned} f_o(t) &= f_i(t - \tau) = \\ &= f_i(t) - \tau f'_i(t) + (\tau^2/2) f''_i(t) + \dots \end{aligned} \quad (4)$$

It becomes nonlinear if τ depends on f in any way, for example if $\tau = b f_i(t)$. Then, the signal transmission function is

$$f_o(t) = f_i(t) - b f_i(t) f'_i(t) + (b^2/2) f_i^2(t) f''_i(t) + \dots, \quad (5)$$

which may also be written as representing automodulation,

$$f_o(t) = f_i(t) [1 - b f'_i(t) + \dots], \quad (6)$$

again indicating a possible measure of the departure from linearity by an automodulation index, the product of b with

some measure of the magnitude of f'_i . Early successful analyses of tracing error were based on such expansions.^{2,3}

Further analyses have the goal of developing well defined indices of distortion measuring the degree to which the distorted signals depart from their desired quality. For auditory signals, because of the pitch analysing properties of human hearing, indices measuring alterations of the spectral properties of the transmitted signal have a special aptness, so that further analysis is Fourier analysis. In general, there are two ways in which the spectrum may be distorted. For one, the strengths and phases of spectral components already present in the original signal may be altered. While the ear is sensitive to such alterations, especially of strengths of components, those changes are useless as indices of nonlinear distortion, since those very changes may also be produced in linear systems. The distortion is called frequency distortion and is of no concern here.

The second way in which the spectrum may be altered is by the generation of new spectral components not present in the original signal. Such alteration is the hallmark of nonlinear distortion. It stems directly from the frequency translating or heterodyning properties of the modulation process. The new frequencies will be the harmonics of all those originally present, together with all frequencies that may be calculated as sums and differences between those originally present and between those and the new harmonics. In judging the audibility of these spectral alterations, one has recourse to the empirical data on masking.⁴ These data make such a judgment possible for only a very few tones simultaneously present. Also, it is the relative strength of those tones that counts, without regard for phase.

For the purpose of planning objective experimental measurements, it is important to develop more detail than is required for the application of masking theory. A need to correlate the measurements with particular nonlinear mechanisms makes phase information as valuable a classification tool as are frequency and amplitude. Modulation concepts are also helpful in classifying spectral components into groups, and in identifying the responsible mechanisms.

Some of the components may be identified with cross-modulation. If a linear combination of signals, $f_{11} + f_{12}$, be subjected to the nonlinearity expressed by Eq. (3) there result, in addition to the automodulation terms ($f_{11}^2 + f_{12}^2$), the crossmodulation terms

$$f_{11}(1+af_{12}), \quad f_{12}(1+af_{11}).$$

Similar terms are called crossmodulation when the two signals are transmitted in channels which are physically distinct but are allowed to interact to produce such terms in the absence of automodulation. Crossmodulation should not be confused with crosstalk, which describes a simple additive or linear contamination of the signal in one channel by that in the other.

The two channels may not be physically distinct, but may involve transmission in a common medium with the channels being distinguishable because their two spectra involve wide separation from one another in frequency. Again, the crossmodulation concept is useful. It is, of

course, possible to speak of separate parts of the spectrum belonging to only one signal. Then crossmodulation between these parts is more properly termed intermodulation.

Apart from an identification based on which signals are involved, there is an identification based on the structure of the modulation. In this way, amplitude modulation may be distinguished from phase modulation. This is possible only for narrowband signals. That is, the representation

$$f(t) = A(t) \cos[\omega_0 t + \psi(t)], \quad (7)$$

is unambiguous for the identification of $A(t)$ as representing amplitude modulation and $\psi(t)$ as representing phase modulation only for the narrowband case.⁵ A narrowband signal is one for which the spectral components of significant strength occupy a frequency interval which is small compared to the frequency for such a component. Strictly speaking, the frequency band should be vanishingly small compared to any frequency in the band, but such exactitude is beyond the needs of the present classification scheme. Signals which are not narrowband are called baseband.

As an illustration of this important theorem, consider subjecting a simple sinusoidal test signal to nonlinearities such as those expressed in Eqs. (3) or (6). The spectrum of the output signal is baseband, because only harmonics

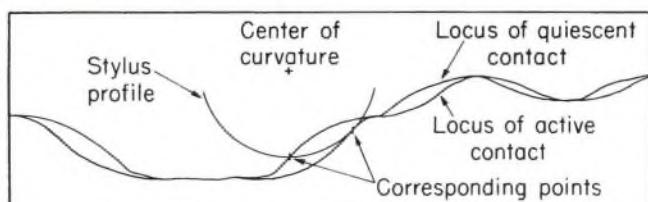


FIG. 3. Waveforms connected by the tracing transformation. The two curves shown, one always lying below the other, are related by both a delay and an amplitude perturbation. Since the slopes of the two waveforms must agree at corresponding points, the delay perturbation is the more fundamental, there being no amplitude perturbation in the corresponding slope waveforms. In the absence of a systematic marking of corresponding points or of information showing the construction of one waveform from the other, the difference between the curves could as easily be ascribed to only an amplitude perturbation as not. The ambiguity is inherent in the lack of any systematic fine structure in the waveforms themselves.

are involved. It is then pointless to examine this spectrum to determine whether the distortion arose because of an amplitude-perturbing mechanism, with or without memory, as in Eq. (3) or its extensions, or whether it arose because of a delay-perturbing mechanism, as in Eq. (6) or its extensions. The same sort of ambiguity is evident in the pair of waveforms exhibited in Fig. 3. These are curvature-limited, but otherwise quite arbitrary, baseband waveforms connected by the tracing transformation. It has been shown that the transformation involves delay perturbation⁶ and nothing more,^{7,8} but there is no characteristic in these waveforms that requires such an interpretation or even particularly suggests it.

The trouble is that the waveforms of Fig. 3 are too smooth and featureless, that there is no finely detailed regularly repeating structure in one that may be used as a basis for establishing correspondences with a transformed fine

structure in the other. Suppose, for example, the draftsman had placed tic marks at equal intervals along the lower curve and had marked the corresponding transformed points on the upper curve. If those tic marks had turned out to be equally spaced as well, the hypothesis of a delay-perturbation mechanism would have to be ruled out. Otherwise, the delay variation would be plainly evident, and the finer the interval that had been chosen, the more accurately could the amplitude and delay variations between the marks be interpolated. This is, of course, the qualitative substance of the quoted narrowband theorem. The oscillations of the cosine of Eq. (7) are to mark the intervals which would be equally spaced if $\psi(t)$ were constant. The more slowly $\psi(t)$ varies, in comparison to this oscillation, the more narrowband the waveform is, and the more oscillations there are between significant changes in $\psi(t)$. Similarly, the hypothesis requires $A(t)$ to be slowly varying on the same time scale.

Thus it is that an ability to discriminate between delay-perturbing and amplitude-perturbing nonlinearities by instrumental analysis of the resulting waveform requires that a meaningful part of the spectrum be used which is narrowband. In the SMPTE intermodulation test, the two test signals have frequencies separated by a large ratio. Frequencies showing a 10:1 ratio, such as 400 and 4000 cps, presented at a 4:1 amplitude ratio, constitute a popular version, though larger ratios, as in 60 and 6000 cps, are sometimes used. The higher frequency signal is identified as the "carrier," and the test for nonlinearity consists in the observation of the modulation of the carrier on the part of the lower frequency tone. Only those Fourier components in a narrow band of frequencies centered on that of the carrier need be presented to the modulation detector. Thus, this is the test allowing discrimination between amplitude- and delay-perturbing nonlinearities; it predominantly tests the nonlinearity acting on the stronger lower frequency tone.

Another "intermodulation test" uses tones arrayed in the CCIF configuration. Here, the test signals themselves constitute a narrowband signal, there being two tones whose separation in frequency is small compared to the frequency of either. A popular combination is 4000 and 4400 cps tones of equal strength. Neither tone may be identified as a carrier, and the test for nonlinearity consists in the observation of a low-frequency "beat" tone, such as 400 cps, or in the observation of other sum or difference tones. The spectrum being observed is, in any case, a baseband one and thus one for which the concept of modulation has no direct meaning. Thus, the test shares with that for harmonic distortion an inherent failure to provide a direct discrimination between amplitude and delay perturbing nonlinearities.

AMPLITUDE PERTURBATIONS

The wide variations in indicated angles for minimum tracking error shown in Figs. 1 and 2 are caused by small amounts of extraneous amplitude-perturbing nonlinearities present as a contamination. There are two reasons for the fact that small contaminations produce such serious errors. The first is that the measures based on amplitude inter-

modulation and harmonic distortion are sensitive to amplitude perturbations, and the second is that these same measures are only weak indicators of delay perturbations.

Though many factors may account for the amplitude perturbing nonlinearities, and a fully adequate explanation would require each to be considered, only one will be selected for illustration. It will be seen that it suffices to account for nearly all of the major anomalies in the measurements. It is a tribute to the care with which the measurements were made that they do not show even graver anomalies.

From Eq. (6) by a suitable identification of b , or from more complete analyses (See Ref. 8 for details and a bibliography), one may calculate the second harmonic content due to tracking error. For a sinusoidal displacement signal, recorded with effective cutter inclination K , the signal being played back with a pickup inclination κ , the displacement signal is

$$\sin \omega t + \frac{1}{2}(K-\kappa)(v/v_g)\sin 2\omega t + \dots \quad (8)$$

written out to include the second harmonic, with the assumption that the phase automodulation index

$$\beta = (K-\kappa)(v/v_g) \quad (9)$$

be not too large compared to unity. Here, v is the peak recording velocity, v_g is the groove speed, and "inclination" refers to the tangent of the cutting or tracking angle, rather than to the angle itself.⁸ The distinction is analogous to that between "slope" and "angle of slope," except that inclination here denotes a departure from perpendicular. The velocity signal corresponding to expression (8) would be

$$\cos \omega t + (K-\kappa)(v/v_g)\cos 2\omega t + \dots \quad (10)$$

Suppose that there is some amplitude-perturbing nonlinearity in the cutter driving electronics, and that it acts on the velocity signal. It will distort the cosine velocity signal, as may be seen from Eq. (3), to

$$\cos \omega t + (av_g/2)\cos 2\omega t + \dots \quad (11)$$

The combination of this with tracking error leads to the distorted wave

$$\cos \omega t + (K-\kappa+av_g/2)(v/v_g)\cos 2\omega t + \dots \quad (12)$$

Thus, it is seen that the minimum in second harmonic distortion would come not for $K=\kappa$, but for a pickup inclination of

$$\kappa = K + av_g/2. \quad (13)$$

If from Fig. 1 the cutting angle is taken to be correctly indicated from the FIM curve as 18.8° , then K would be 0.34. However, the second harmonic distortion is minimum at 13.5° , or a κ of 0.24. There is calculated $av_g = -0.20$. A negative value for this coefficient is typical of saturation types of amplitude perturbations (limiting). The groove speed corresponding to a 3.90 in. radius is 13.6 ips or 34.5 cm/sec. Thus the coefficient is $a = -0.58$ percent per cm/sec of recording velocity. The recording level was for a peak velocity of 7 cm/sec, so that the second harmonic con-

tent due to the amplitude perturbing nonlinearity must have been $a v^2 = 2.0$ percent in the velocity waveform, or only one percent in the corresponding displacement waveform. Tolerance of this much distortion could hardly be said to represent bad practice at such recording levels, yet the disturbance in angular measurement is appreciable.

The disturbance is reminiscent of that used by CBS Laboratories in generating synthetic cutting angles, as recently described by B. B. Bauer,¹⁰ for the production of test record STR-160. The calibration of the record automatically includes the effect of all harmonic terms capable of affecting the angle indication, presumably, but the user takes upon himself the responsibility for assuring that no distortion terms capable of spoiling that calibration are introduced during playback. This is not a trivial matter, in view of the insensitivity of the test.

A similar analysis may be given for the disturbance in indicated angle arising for the amplitude-intermodulation measurement (AIM) in the velocity waveform. The disturbance is numerically equal to that for second harmonic distortion. It turns out, however, that the disturbance for AIM on a displacement basis $|AIM(D)|$ is much greater. The dependence of these disturbances upon recording radius is easily seen via the dependence on v_g . Using a value of a calculated for the disturbance of the AIM velocity-based measurement at the 3.90 in. radius, calculations of the disturbance for the AIM(D) measurement at that radius, and for both at the radii of 2.50 and 5.50 in. have been made.

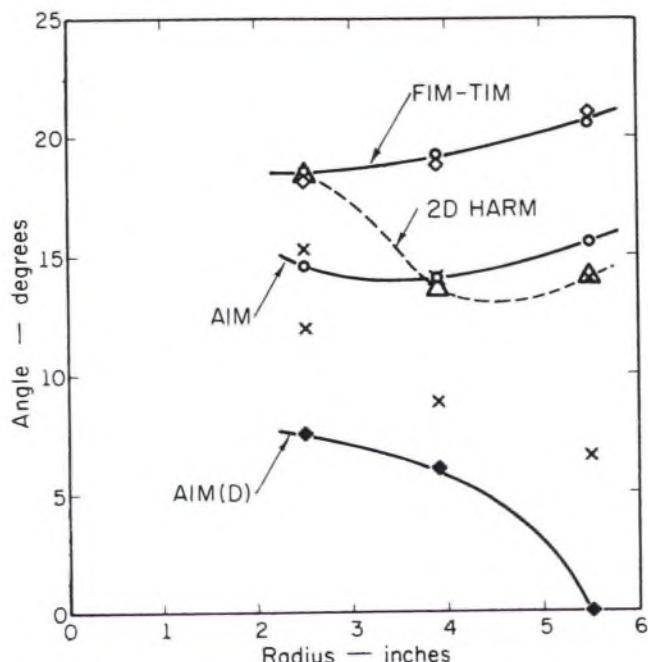


FIG. 4. Theoretical disturbance in angle data. The figure is the same as Fig. 2 except for the points plotted as \times . These points show the disturbance in the angle data to be expected for a small amount of contaminating amplitude-perturbing nonlinearity. The extent of the nonlinearity was estimated from the AIM disturbance at the 3.9 in. radius. For the kind of contamination assumed, the disturbance in the second harmonic data should agree with that for the AIM data. The disturbance in the AIM(D) data should be more severe. A complete explanation of all the anomalies is not obtained.

The results are plotted using the symbol \times in Fig. 4, which is otherwise identical with Fig. 2.

The numerical agreement must be considered satisfactory in view of the fact that only one of the many possibilities for the source of the disturbing distortion has been assumed. It would be presumptuous, on the basis of the data presented, to seek to identify the other sources. It is true, however, that the AIM(D) for tracking error is a measure related not to the effect at 400 cycles, as it is for the others, but to the effect at 4000 cycles.¹¹ As will be seen, the effect at 4000 cycles may be quite different, including the possibility for phase shifts tending to simulate tracing distortion.

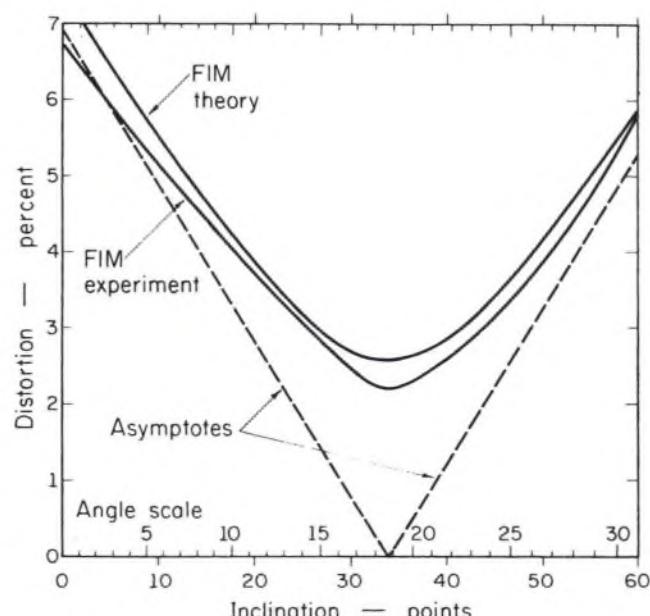


FIG. 5. Comparison of theoretical and experimental curves for frequency intermodulation distortion. The experimental data are those of Fig. 1. The theoretical curve is calculated assuming the given conditions for the tracking error contribution and a 1.0 mil stylus radius for the tracing error contribution. No other perturbations were assumed. The theoretical curve is hyperbolic on a linear inclination scale (tangent scale) whose asymptotes would give the tracking distortion in the absence of tracing error. The inclination scale is such that the tangent of 45° is 100 points. An angle scale is shown for reference.

These possibilities could account for the extra disturbance of that measurement, in part, and for the shallow character of the minimum.

Incomplete as the present explanation may be, it is sufficient to condemn second harmonic distortion or amplitude intermodulation distortion as not apt or consistent indices of tracking error, and not fit for use as a standard. On similar grounds, they should be dismissed as measures of tracing error. It remains to be seen whether the FIM measurement is a satisfactory index.

Certainly, the use of a linear pulse-count discriminator (for example General Radio Frequency Meter and Discriminator Type 1142-A) should provide abundant immunity from the effects of spurious amplitude perturbations. Since this was the instrument used,¹ there is a strong *a priori* presumption of validity. Whether there can be delay-

perturbing nonlinearities other than tracking or tracing error to introduce spurious indications is a question examined in a later section. For the present, tests of internal consistency may be applied.

One of these is the variation of indicated cutting angle with groove radius. The variation is slight enough to allow the judgement that it is a real feature of the cutting process. Another test involves the shape of the FIM curve shown in Fig. 1. As stated, the shape should be hyperbolic, because of the root-sum-of-squares rule for combining the distortion indices of tracing and tracking error.⁸ Actually, the shape should be strictly hyperbolic only if the angle scale is replaced by an inclination scale in which the angle is replaced by its tangent. The units on this scale are called "points" and have been chosen, for convenience, so that an angle of 45° corresponds to 100 points of inclination.

In Figure 5 the FIM data are shown plotted on a uniform inclination scale. Also shown, for reference, is a scale of angles; that scale is only slightly non-uniform. The curve may be compared with the theoretical curve which has been adjusted to show a minimum at the same inclination, 34 points or 18.8 degrees. The value at the minimum is the theoretical value for the FIM measure of tracing distortion, for the given conditions, assuming a stylus radius of 1.0 mil. The theoretical curve is a hyperbola whose asymptotes are the straight lines shown. The slopes of these asymptotes are the theoretically correct ones for the tracking distortion. The fit is excellent, if one makes allowance for the fact that

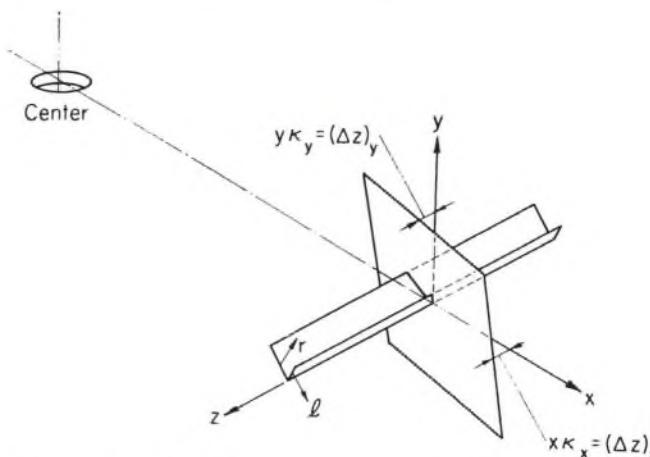


FIG. 6. Geometry of the stereo groove and plane of stylus motion. The z -axis points in the direction of record motion. The x and y axes define positive directions for lateral and vertical displacements respectively. The positive directions for l and r displacements are also shown. The stylus plane is shown inclined with respect to the x,y plane, with the amount of inclination defined in terms of an x -component and a y -component. The orientation of axes with respect to the center hole is shown.

the original data had been plotted to semilogarithmic scales and that the published curves involved a judicious interpolation, in all probability using French curves. Thus any observed anomalies are probably draftsman's artifacts.

Since the FIM data pass this consistency test with flying colors, it is necessary to examine the possibilities for other delay-perturbing nonlinearities that could produce system-

atic errors in the measurements. In the process it will be shown that a measurement showing a much deeper null in the presence of tracing error may be devised.

CO-CHANNEL DELAY PERTURBATIONS

Co-channel delay perturbation describes a nonlinear process in which two channels suffer a common delay perturbation, controlled by a combination of the signals present in the two channels. The prime example is the perturbation caused by tracking error in reproduction of the signals in the two-channel stereo groove. The stylus motion necessarily has a component in the direction of the groove axis, if the stylus is constrained to move in a surface not containing the normals to the groove walls. Further, it must move as a single unit, so that its axial motion produces delay perturbations equally for both channels. The result is a combination of crossmodulation as well as automodulation and intermodulation depending on exactly which combination of signals excites the axial motion.

The details of the two-channel tracking-error situation may be worked out with the help of Fig. 6 which shows the geometry of the stereo groove. Rectangular x,y,z coordinates are shown with origin at the groove axis for a quiescent groove. The x -axis lies in the plane of the disc and points away from the center hole. The y -axis is normal to the disc and points upwards. The z -axis lies along the groove axis and points in the direction of record motion. The positive directions for the displacements of the groove walls are indicated. They involve combinations of vertical and lateral displacements $(x+y)/\sqrt{2}$ for the right channel and $(x-y)/\sqrt{2}$ for the left.

Also shown in Fig. 6 is a plane tipped away from the x,y plane about some arbitrary axis of rotation. This plane is to be thought of as the one containing the stylus motion. More precisely, though the distinction will not be needed since the motions are always so small, it is the plane tangent to the surface containing the stylus motion. The stylus plane intersects the y,z plane in the line whose equation is

$$(\Delta z)_y = \kappa_y y,$$

and the z,x plane in the line whose equation is

$$(\Delta z)_x = \kappa_x x.$$

Thus it is readily seen that the equation of the plane[†] is

$$\Delta z = \kappa_x x + \kappa_y y. \quad (14)$$

The interpretation is that for x,y as the components of a displacement vector, κ_x, κ_y are the components of an inclination vector in the x,y plane. The invariance of the stylus plane when described in coordinates rotated in any way about the z axis shows that the scalar product of Eq. (14) is a true scalar, and that the inclination is a true vector of invariant length $\sqrt{(\kappa_x^2 + \kappa_y^2)}$. This "length" is the total

[†] There is a related homogeneous equation for the stylus plane in terms of the plane's unit normal vector. This vector has three components, which are the familiar direction cosines. The present representation, giving a special role to the z -axis and using only a two-component vector, is more natural in the discussion of tracking error.

inclination and it is the tangent of the angle between the stylus plane and the x,y plane. The direction of the vector is normal to the intersection of these two planes. For example, the vector points in the direction of the y axis for a clockwise rotation of the stylus plane viewed from the center hole.

These observations guarantee that it always makes sense to speak of the inclination being resolved along other orthogonal axes in the x,y plane. Thus, one may speak of a right and left component for the inclination of the stylus plane. In terms of those components the stylus plane would be

$$\Delta z = \kappa_r r + \kappa_l l, \quad (15)$$

where r and l refer to the displacements for the right and left channels. This is an example of the invariance of the stylus plane guaranteeing the proper transformation of the components of the inclination vector. It has been shown that the components are additive for compounding skew transformations⁸ so that the same rule must apply to the inclination vector as a whole. Thus the vector difference between the effective cutting inclination and the tracking inclination is the tracking error vector having components

$$\Delta \kappa_r = \kappa_r - K_r, \quad \Delta \kappa_u = \kappa_u - K_u \quad (16)$$

where the capital letter refers to the cutting inclination.

The same tracking error vector has the components

$$\begin{aligned} \Delta \kappa_r &= (\Delta \kappa_x + \Delta \kappa_y) \sqrt{2}, \\ \Delta \kappa_l &= (\Delta \kappa_x - \Delta \kappa_y) \sqrt{2}. \end{aligned} \quad (17)$$

referred to the r,l axes. The total delay perturbation is

$$\begin{aligned} \Delta z &= \Delta \kappa_r r + \Delta \kappa_l l, \\ &= \Delta \kappa_x (r+l) \sqrt{2} + \Delta \kappa_y (r-l) \sqrt{2}. \end{aligned} \quad (18)$$

showing the linear combinations of the two-channel displacements accounting for the delay perturbation common to both channels. The second of equations (18) follows from Eq. (17) by substitution. Since $(r+l) \sqrt{2}$ is the x displacement and $(r-l) \sqrt{2}$ the y one, the invariance of the scalar product, regardless of basis, is also illustrated. The single-channel distortion products and the two-channel cross-modulation products have already been calculated.^{7,10}

These motions are governed by geometrical factors alone (kinematics); motions governed by dynamics are more complicated to specify, since they involve motions of bodies governed not only by geometry but also by the mechanical, elastic, viscous, and inertial reactions in response to driving forces related to the signal variables. The reactions may be nonlinear functions of the signal variables, beyond the nonlinearity inherent in the generation of co-channel delay perturbations. These perturbations may be expected to show the strong frequency dependence characteristic of mechanical reactances. For the stylus-groove interaction, very little is known about such processes, so that a sketch, rather than an adequate survey, must suffice.

Of the dynamic processes producing co-channel delay perturbations the best understood is recording stylus springback.¹¹⁻¹³ The driving force appears to be proportional to the square of the instantaneous depth of cut and very little

dependent on groove speed over the range of interest. The mechanical impedance controlling the axial motion appears to be dominated by a nearly linear stiffness at low frequencies. Thus for small deviations from the mean depth of cut, the instantaneous delay perturbation is in phase with, and proportional to, the instantaneous displacement signal. The net effect is a simple reduction in the vertical cutting angle.

At higher frequencies, the effect will change as the mechanical impedance will begin to show the effect of inertial reactance. As this begins to happen, a viscous component in the impedance will dominate in the neighborhood of the flexural resonance for axial motions of the stylus tip, so that the axial motions will be in quadrature with the displacement signal. The distortion components will then resemble those for tracing distortion. At still higher frequencies inertial reactance will dominate, causing the cutting angle to increase. The frequency for the resonance is not known but may be estimated to come within the audio band, possibly within a factor two of 5000 cps.

Less is known about lacquer springback.¹¹⁻¹³ It clearly contributes to a reduction in effective cutting angle by an observable amount but to an extent comparable to, or less than, the extent attributable to stylus springback. Its frequency characteristics are unknown. If the dynamic system is one for the propagation of acoustic waves, the thought that these propagate more rapidly than the groove speed suggests that there should be no frequency-dependent anomaly.

Recently Bauer has called attention to playback stylus springback¹¹ as explaining observed tracking angles greater than predicted from pickup geometry. Spurious axial motions of this kind were observed by Rabinow and Cordier¹⁵ and analyzed by Bayliff.¹⁴ The observations by Rabinow and Cordier are obsolete, and Bayliff made no conclusive observations. The driving force is unknown, though Bayliff made detailed analyses based on the thought that the motion is driven by harmonics of the displacement signal. He showed the effects of various locations for the frequency of the resonance for axial motion of the stylus. These frequencies are not known for commercially available pickups. Potential complications from this source should be suppressed by designing pivots to show great axial stiffness.

A test designed to discriminate against amplitude-perturbing nonlinearities, such as the FIM test of Halter and Woodward,¹ will nevertheless have to contend with these co-channel dynamic effects. Hopefully, they may all be expected to be simple at low frequencies, leading only to a shift in the cutting or tracking angle. Therefore, it is useful to invoke the concept of effective angle, at least for the low frequencies near 400 cycles. Then, if these angles may be expected to stay reasonably fixed, they can, with suitable calibration, be standardized. The manner in which these effects may be dealt with at higher frequencies must wait upon further study. In any case, it is clear that the standardization must be based on dynamic tests realistically simulating practices in the field but invoking laboratory precision. Static calibration, based on geometrical measure-

ments, will be seen in a later section to be useful only for cartridges designed, as test instruments, to show mechanical stability of an order not necessarily required of high performance reproducers.

There remains a second class of delay-perturbing nonlinearities which will be examined in the next section. As will be seen, sensitivity to these may be expected to be reduced in a crossmodulation test.

IN-CHANNEL DELAY PERTURBATIONS

In-channel delay perturbation describes a nonlinear process in which each channel suffers a delay perturbation controlled by the signal in that channel alone. Corrington and Murakami¹⁶ observed that tracing distortion arises independently in each of the *r,l* channels of the stereo groove. In Fig. 7 it is seen that the rounded stylus tip,

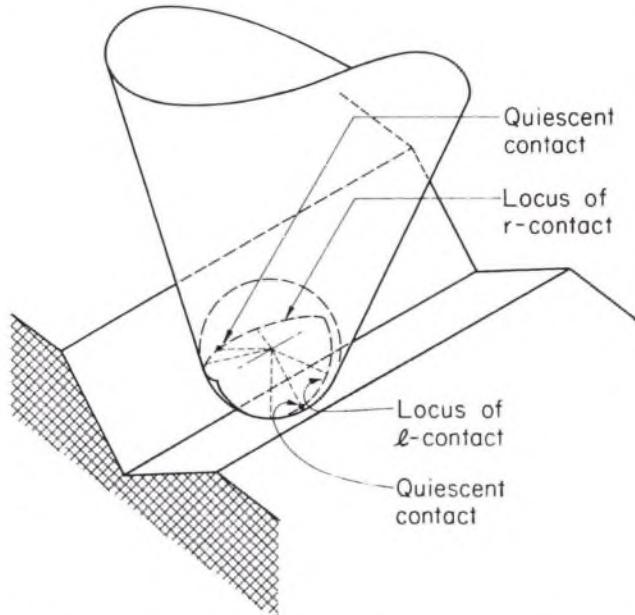


FIG. 7. Geometry of the stereo groove and surface of stylus tip. The stylus is represented as conical with the tip rounded to a spherical shape. The two "points" of contact with the two groove walls are shown for an unmodulated (quiescent) groove. When modulated, the groove walls will show slopes causing the points of contact to move to new positions along the loci of contact, independently for each groove wall. In this way, each channel suffers its own delay perturbation. The loci of contact must lie in planes normal to the groove walls but parallel to the quiescent groove axis.

normally somewhat spherical in form, contacts each groove wall separately. Then, if each groove wall be separately modulated by displacements normal to it, the slope of the displacement will cause each contact point to shift relative to its quiescent position on the stylus. The shift of the contact for a given wall will be independent of the slope of the modulation for the other wall.

It is just this shift in the contact point which is the delay perturbation for the channel in question. Further, it has been shown that the whole of tracing distortion is generated by this delay perturbation of the slope waveform.^{7,8} These observations establish tracing error as a prime example of in-channel delay perturbation.

Treatment of the two-channel properties of in-channel delay perturbation requires the calculation of distortion components for each of the *r,l* channels separately. Then, if it is desired to consider other pairs of channels, such as the vertical-lateral pair, these components may be resolved into those channels and the resulting combinations studied. Due attention must be paid to signs so that cancellation effects may be noted.

These signs may be traced using an extension of an analysis already given in an earlier article.⁷ As noted there, the Fourier analysis of the perturbed waveform may be given in closed form if it is assumed that the locus of contact with the groove wall is parabolic. This hypothesis was noted to be as realistic as any for practical styli and will be used here. The tracing of signs will be correct, however, for any stylus shapes showing the same symmetry for the loci as the even symmetry of the parabola.

If the unperturbed waveform for the *r*-wall is taken to have the slope described by

$$s_r(z) = (v_r/v_g)(\epsilon_{1r} \sin k_1 z + \epsilon_{2r} \sin k_2 z), \quad (19)$$

then the cited analysis⁷ gives the Fourier component for the frequency

$$k_{mn} = mk_2 + nk_1$$

to be, for the perturbed slope waveform

$$S_{rnm} = 2(v_r/v_g)J_m(\epsilon_{2r}\beta_{mn})J_n(\epsilon_{1r}\beta_{mn})/\beta_{mn}, \quad (20)$$

where

$$\begin{aligned} \beta_{mn} &= m\beta_2 + n\beta_1 \\ \beta_2 &= c(v_r/v_g)k_2 \\ \beta_1 &= c(v_r/v_g)k_1 \end{aligned}$$

and v_r is the peak velocity, v_g is the groove speed, k is the radian wave number, i.e.,

$$k = 2\pi/\lambda = 2\pi v_r, \quad v_g = \omega_r v_g,$$

where λ is the wavelength, and v is the frequency in cps. The parameter c is the radius of curvature of the contact locus at the quiescent contact point. In Eq. (20), $J_n(\cdot)$ stands for the Bessel function of first kind, order n .

Very little knowledge of the properties of Bessel functions is required for their use in these formulae. For example, numerical values of usually sufficient accuracy may be obtained by using

$$J_n(\beta) \approx (1/n!)(\beta/2)^n [1 - (\beta/2)^2(n+1)! + \dots], \quad (23)$$

for n positive and β not larger than 1. Often the bracketed expression may be replaced by unity. For negative values of n or β , one needs the symmetry relations

$$J_n(-\beta) = (-1)^n J_n(\beta), \quad (24)$$

$$J_{-n}(\beta) = (-1)^n J_n(\beta),$$

which will be used in tracing signs.

If the form of Eq. (19) is also used, for $s_l(z)$ for the unperturbed slope waveform for the *l* channel, one gets the same components as in Eq. (20) except that the subscript

r is replaced by l throughout, and there is also the multiplying factor

$$P_{mn} = -(-1)^{m+n}. \quad (25)$$

The reason is that the direction of positive displacement for the l channel is defined in the opposite sense, with respect to the outward normal, from that for the r channel. The effect is to reverse the sense in which the sign of the stylus curvature is to be understood, which reverses the sign of β_{mn} . Use of Eq. (24) gives the factor shown as Eq. (25).

The parameters ϵ_1 and ϵ_2 can be used to adjust the relative magnitudes and signs of the two unperturbed waveforms for investigating certain combinations. For example, one may set $\epsilon_1 = 1$, $\epsilon_2 = 0$, for both channels. The representation is then for a single tone recorded laterally. Then in Eq. (20) one must set $m = 0$, $n \geq 1$. Further, one may take $v_r = v_l = \frac{1}{2}v\sqrt{2}$. For S_{rn} one has

$$S_{rn} = 2(v_r/v_y)J_n(n\beta_1)/n\beta_1, \quad (26)$$

and the same for S_{ln} , except for the factor $-(-1)^n$. One may express β_1 as

$$\beta_1 = \frac{1}{2}\sqrt{2c(v/v_y)k_1},$$

by replacing v_r or v_l by its value $\frac{1}{2}v\sqrt{2}$. As Walton observed¹⁷ this is equivalent to using a reduced radius for the stylus. Writing

$$S_{rn} = \frac{1}{2}\sqrt{2}(S_{rv} + S_{lv}), \quad S_{ln} = \frac{1}{2}\sqrt{2}(S_{rv} - S_{lv}), \quad (27)$$

one obtains, finally

$$S_{rn} = 2(v/v_y)P_{rn}J_n(n\beta_1)/n\beta_1. \quad (28)$$

where

$$P_{rn} = \frac{1}{2}|1 - (-1)^n|. \quad (29)$$

The formula for S_{ln} is the same except that P_{rn} is to be replaced by

$$P_{ln} = \frac{1}{2}|1 + (-1)^n|. \quad (30)$$

These indicators P_{rn} and P_{ln} show how the components are distributed among the vertical and lateral channels. They place the odd harmonics, for $n = 1, 3, 5, \dots$, in the lateral channel and the even ones for $n = 2, 4, 6, \dots$, in the vertical. This sorting describes the well known "pinch effect" for lateral recording.

Specialization to vertical recording may be made by reversing the sign of ϵ_{1r} . The formulae are all the same except that the sign factor $-(-1)^n$ becomes $-(-1)^n(-1)^n = -(-1)^{2n} = -1$. The result is that $P_{rn} = 0$ and $P_{ln} = 1$, that is, all the components lie in the vertical channel.

Two tone lateral recording is similarly treated. One sets $\epsilon_{1l} = \epsilon_{1r}$ and $\epsilon_{2l} = \epsilon_{2r}$. The results closely parallel those given, except that the indicators are

$$P_{cmn} = \frac{1}{2}[1 - (-1)^{m+n}], \quad (31)$$

$$P_{ymn} = \frac{1}{2}[1 + (-1)^{m+n}].$$

These place the odd harmonics of each tone in the lateral

channel and the even harmonics of each in the vertical. For the sum and difference tones, the odd-even and even-odd ones

$$m = 1, 3, 5, \dots; n = \pm 2, \pm 4, \dots,$$

$$m = 2, 4, 6, \dots; n = \pm 1, \pm 3, \dots,$$

are in the lateral channel, but the even-even and odd-odd ones are in the vertical.

For both tones recorded in the vertical channel, all components of the perturbed waveform appear in the vertical. All of these calculations are based, of course, on the hypothesis that the stylus curvature is the same in magnitude for both groove walls. Otherwise the cancellations of components would be only partial, and the formulae would be more complicated.

Although tracing error cannot cause crossmodulation between r and l channels, it readily does so for the vertical and lateral channels. To see this one selects $\epsilon_{1r} = -\epsilon_{1l}$ and $\epsilon_{2r} = \epsilon_{2l}$. Then a signal of frequency ω_1 is recorded vertically, but ω_2 is recorded laterally. The distortion components become distributed according to

$$P_{xmn} = \frac{1}{2}[1 - (-1)^{m+2n}] = \frac{1}{2}[1 - (-1)^m], \quad (32)$$

$$P_{ymn} = \frac{1}{2}[1 + (-1)^{m+2n}] = \frac{1}{2}[1 + (-1)^m].$$

In this way, one finds all harmonics of ω_1 in the vertical but the harmonics of ω_2 divided: even in the vertical and odd in the lateral. The sum and difference tones for the vertical are for

$$m = 2, 4, 6, \dots; n = \pm 1, \pm 2, \pm 3, \dots,$$

but in the lateral they are

$$m = 1, 3, 5, \dots; n = \pm 1, \pm 2, \pm 3, \dots.$$

Of these the important 1.1 sum and difference tones are in the lateral channel. Interchanging the roles of ω_1 and ω_2 still leaves these 1.1 tones in the lateral channel.

The adequacy of using formula (23) for numerical work has been discussed elsewhere,¹⁸ but some numerical examples will be instructive. The tracing conditions of Fig. 1 lead to the estimates $\beta_1 = 0.0262$ and $\beta_2 = 0.262$. Using $J_0(\beta) = 1$ and $J_1(\beta) = \beta/2$ as reasonable approximations, the first column of Table I is calculated. The second column shows more exact values calculated from tables of Bessel functions. The experimental values, read from Fig. 1, are entered in the third column. There the entry TIM is simply the sum of the first order tones, without regard to sign, expressed as a percentage of the 4000 cycle tone present. The letters stand for total intermodulation. The amplitude intermodulation figures are derived in a similar way from the difference in those sidetone magnitudes. The approximate and exact results show excellent agreement in every particular.

Agreement with the experimental values, while reasonable, may be largely fortuitous. Other system distortions can disturb the measurement, as they do for tracking distortion. Also, since the measurement does not discriminate against co-channel perturbations, the contribution to this quadrature component by a quadrature component in the cutter

springback motion, described in the preceding section, may not be ruled out. Already, the need to choose a 1.0 mil stylus, to fit the quadrature component in the FIM data, is a suspicious circumstance, since, though it is nowhere mentioned in Ref. 1, a 0.7 mil stylus was probably used. The TIM and second harmonic values are more nearly consistent with the latter radius. Whether the anomalously large AIM(D) value may be attributed to "other system distortions," quadrature springback, or groovewall deformation is a determination which also must await further study.

The observation that tracing distortion is a little *less* than calculated has been attributed to the in-channel perturbation, for amplitudes at least, resulting from a groove-wall deformation. The effect is still imperfectly understood. There is some possibility for a delay perturbation. If the dynamic contact force is normal to the wall, the deformation will be in the same direction. However, the instantaneous wall normal has an axial component. More than this, sliding friction generates instantaneous contact forces which place the stylus-groove-wall interface under traction, generating deformations which could have axial components as large as the normal components. Mathematical tools for dealing with contact stresses involving interface traction have only recently become available¹⁸ and have not yet been applied to the stylus-groove interaction.

The in-channel character of these deformations holds so long as the stylus admittance is isotropic in the surface of stylus motion. Walton has observed deformations in both walls resulting from a signal impressed on one wall only.¹⁷ The frequency was such that the admittance was inertia dominated. It must be supposed that the inertia was not isotropic: such would be the case only for bodies constrained to rotate about pivots. Walton's mechanical system is evidently not so simply described. In any case, such a mechanical cross-coupling can introduce co-channel effects into what is ordinarily an in-channel perturbation, especially at frequencies near the stylus-groovewall resonance.

The most violent of all in-channel perturbations can be the loss of groovewall contact, resulting from the application of insufficient static bearing force. The distortion is easily heard, and it may be the most often heard in poorly maintained systems. Despite its probable prevalence, it is so easily corrected that it is customarily dismissed with little comment in discussions like the present one. That custom is followed here.

STANDARD TRACKING ERROR TEST

In the light of all of these theoretical considerations, it is seen that a tracking error test should discriminate between amplitude perturbations and delay perturbations to avoid systematic errors. It should also discriminate between co-channel and in-channel delay perturbations, so that the accuracy of the desired measurement (co-channel) need not depend on the consistency or stability of the undesired (in-channel) one. Finally, the stability of the tracking error test itself is subject to possible anomalies at high frequencies, so that only a low-frequency (400 cps) test should be proposed as a standard one.

These observations force the abandonment of harmonic distortion and amplitude intermodulation tests, and require the use of a test indicative of phase modulation. Such a test must use test tones in an SMPTE configuration so that the requisite narrowband waveform may be obtained: the CCIF configuration should not be used. Rejection of in-channel perturbations requires the use of a crossmodulation test between the *r* and *l* channels.

In this way a tracking error test suitable as a standard is almost uniquely specified. There remain some parameters that may be selected as a matter of convenience. For example, there is nothing sacred about 400 cps, except that it seems to be a suitably low frequency for which much

TABLE I. Minimum distortion values for the conditions of Fig. 1.

Mode	Tracing theory		Experimental Observed (%)
	Approx. (%)	Exact (%)	
TIM	26.20	26.216	19 *
FIM	2.620	2.620	2.2
AIM	2.620	2.575	2.8
AIM(D)	0.000	0.047	4.1
2nd harmonic	1.310	1.300	0.9

* Not shown in Fig. 1; see Ref. 1.

test equipment is designed. For the higher frequency, 4000 cps, also a popular choice, seems as good as any. Similarly the 4:1 ratio (on a velocity scale) seems as good as any. The actual recorded levels could be 10 cm/sec for the 400 cps signal in one channel and 2.5 cm/sec for the 4000 cps signal in the other, as peak values.

There also remains some choice for the means of observing the modulation of the 4000 cps tone upon playback. The channel in which the 400 cps tone had been recorded will show some side-tones, centered around the missing (because recorded in the other channel) 4000 cps carrier of a phase corresponding to amplitude modulation of that carrier. If these can be sufficiently rejected from the channel in which the 4000 cps carrier had been recorded, and a displacement signal obtained, then there should be so little amplitude modulation that a direct phase modulation observation may be made, using a nearly synchronous phase reference to make a Lissajous pattern on an oscilloscope, as previously proposed.¹⁰ The advantages of this direct PXM method lie in the use of simple widely available equipment for the observation. The disadvantages lie in the difficulty of making accurate readings from a scope, and in the reliance on good channel separation for the AXM rejection.

Total crossmodulation (TXM) observation requires the use of a wavemeter to observe the strengths of the carrier and the upper and lower first-order sidetones. The sum of the sidetones relative to the carrier strength is the measure of TXM. The method is an indirect measure of phase modulation because the phases of the sidetones are not usually measured. Thus, TXM measures phase modulation only if that modulation is larger than the amplitude modulation, otherwise it is the latter that is measured. Calculating the difference in sidetone magnitudes interchanges the two effects. The same procedures for rejecting amplitude

modulation as for the observation of PXM must also be available to TXM to resolve the ambiguity.

The observation of frequency modulation, as reported by Halter and Woodward,¹ is a direct observation equivalent to observing phase modulation. Here, it is frequency cross-modulation (FXM). The use of a linear pulse-count discriminator, as reported, should provide excellent rejection of amplitude modulation. Thus, any mix of the two channels may be used so long as the 4000 cps channel is adequately represented. As in all of these observation schemes, filtering must be used to reject any 400 cps signal present. Otherwise, any equalization mode may be used with FXM observation. There is a side benefit: the use of FXM provides an opportunity to compare the phase of the modulation with the phase of the signal producing it. Unfortunately, the needed equipment, while not expensive is of a rather special-purpose nature for most audio laboratories.

Of the PXM, TXM, and FXM methods of observation, the last-named best suits the needs of a standard and is the choice proposed. The quotation, however, should be standardized to give the peak phase modulation index in percentage units (percent of one radian); this requires that the peak frequency deviation be expressed as a percentage of the modulating frequency (400 cps), or that the peak deviation, expressed as a percentage of the carrier frequency (4000 cps), be multiplied by the ratio of carrier to modulating frequency, a ratio of 10:1. This standardization of the quotation will result in numbers more nearly in agreement with those obtained by PXM and TXM. These observation methods should be regarded as acceptable if not standard observations of tracking error distortion, the former because it requires such simple, readily available equipment, and the latter because of its psychoacoustic significance.

The standard test provides a direct indication of r or l tracking error, depending upon which channel is chosen to contain the low-frequency tone. There is no means of providing a direct test of vertical or lateral error that is not also contaminated** by such in-channel perturbations as tracing error and groovewall deformations. A simple indirect test is available, through providing bands on the test record in which right and left tests alternate. Then, where the two tests give equal results one or the other of the vertical or lateral errors, or both, are null. If the vertical is known not to be null, then the lateral must be so, and *vice versa*.

The null method is not necessary, however. The modulation index may be used to estimate the tracking error rather accurately in the absence of a null. Then, comparing the phase of the 400 cps output from the frequency meter with that of the 400 cps tone on the record, the sign of the error may be determined. Knowing the magnitude and sign for the r and l components allows their calculation for the x

and y components of the tracking error. A convenient laboratory arrangement for adjusting the vertical and lateral errors has been described elsewhere.¹⁹

The slight inconvenience of the r,l test is more than offset by its great accuracy. To check the theoretical accuracy, a calculation of FXM was made for a test recording, using the signals and levels proposed, for a recording radius of 3.9 in. but assuming a channel separation of 26 db for the cutter. Because perfect channel separation was not obtained, some in-channel distortion would appear. This was calculated, assuming it to be tracing distortion for a 1.0 mil playback stylus. The FXM results are plotted in Fig. 8 for

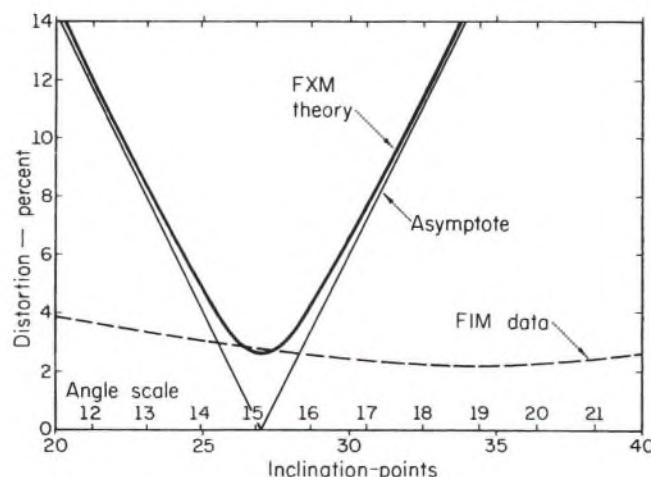


FIG. 8. Tracking error null expected for frequency crossmodulation. The theoretical curve is calculated for a 400 cps signal recorded at 10 cm/sec in one channel and a 4000 cps signal recorded at 2.5 cm/sec in the other, at a radius of 3.9 in. The null was assumed to come at 27 points of inclination (15.2 degrees), for the vertical; the lateral error was assumed to be zero. The null was assumed to be limited by the tracing error resulting from some of the 400 cps signal appearing in the 4000 cps channel as a result of a 26 db cross-talk figure for the cutter, and assuming a 1.0 mil stylus radius. The asymptotes show the tracking FXM for no cutter crosstalk. The FXM figures have been standardized so that the peak frequency deviation is quoted as a percent of the modulating frequency. The nonstandardized data of Ref. 1 is shown as the dashed curve. Taking the standardization into account, the FXM null is seen to be an order of magnitude more sharp and more profound than the FIM null.

varying vertical error. The lateral error was assumed null. A point scale is used for the vertical inclination, and a degree scale is also shown. The FXM values are expressed in the standard way for an FM index (peak frequency deviation as a percentage of the modulation frequency). Also shown are the FIM values (dashed curve) of Halter and Woodward,¹ which had not been standardized. Had the standardization been carried through for those data as well, the two curves would have the same asymptotic slope since the vertical component of the groove modulation was taken to be the same in both instances. Also, the FXM null would have lain a factor of 10 below the FIM null. The present more compact plot shows the FXM null to be more sharply curved by two orders of magnitude, but because of the difference in standardization only one order of magnitude may be claimed. (An alternate plot is presented in a *Letter to the Editor* on p. 344 of this issue.)

** The fact that the 1,1 side tones from tracing error lie in the lateral channel, for vertical-lateral crossmodulation, allows only a lateral tracking error test, with the low-frequency tone in the lateral channel and FXM to be observed for the high-frequency tone in the vertical channel. Even then, the stylus curvatures must be matched at the two groove walls for maximum freedom from contamination by tracing distortion.

The FXM null was assumed to come at 27 points (15.2°) as indicated by the vertex formed by the straight-line asymptotes for the hyperbolic curve. For errors larger than a degree or so, the asymptotes are never more than a quarter of a degree away from the distortion curve. Since these asymptotes would be the distortion curve in the absence of any contamination from in-channel perturbations, tracking errors larger than a degree may be accurately estimated (or over-estimated) by making no allowance for such perturbations. For a two-degree error, the accuracy is one-eighth of a degree. A simple *go-no go* test based on the 4% distortion level would assure conformance to a preassigned standard to within less than a degree.

STANDARD INCLINATION CALIBRATION

The standard tracking error test is capable of revealing an error of about a third of a point of inclination in a total of 27 points. This corresponds to a precision of one part in a hundred for the tangent of a 15° angle or for the angle itself. Calibration to a comparable precision poses a problem not likely to be solved by light-pattern methods. This much precision is available only with the help of an accurately calibrated pickup cartridge.

The static calibration of a specially designed cartridge may be obtained by a microscopic observation of stylus deflections with the cartridge mounted on a precision rotating stage. An orientation is to be sought such that the imposed deflection causes no displacement of the stylus in the direction normal to that deflection. The primary limitation in accuracy stems from an inability to observe a displacement much smaller than a wavelength of light, 0.02 mil. For an applied displacement of 10 mil, the accuracy would be a fifth of a point. A special cartridge design may be needed, because the corresponding requirement on the mechanical stability of the location of the stylus pivot is 0.02 mil for axial motions. The design need offer no particular merit for commercial exploitation, but should allow reproduction by any skilled instrument maker.

In this way a reproducible standard for the calibration of tracking angle may be defined. While it may be unreasonable to expect an industry committee to be responsible for the manufacture of test records or test cartridges, agreement on specifications and the measurement procedures to calibrate instruments may reasonably be expected. Thus, anyone ought to be able to make test records but the tests provided ought to be based on calibrations performed in a standard way with instruments meeting standard specifications.

The proposed test and calibrations offer more precision than is needed for applications in the field. It is not practical, at the present state of the art, to maintain lateral tracking error below a degree or so in the field. The same is true of vertical error, though usually the situation would require tolerance of greater error, especially with changer mechanisms. The establishment of precise laboratory measurement procedures does allow the monitoring of field tolerances with the confidence and the convenience of tests which do not strain the limits of measurability. Also, of course,

precise measurements are an important foundation upon which to build further advances in the state of the art.

TRACING ERROR TESTS

A tracing error test of the same quality as proposed for tracking error does not exist. There is only one test that is inherently uncontaminated by co-channel effects, but it cannot fully discriminate against amplitude perturbations. The tests which do offer such discrimination require one or another of the tracking errors to be set to a null. The latter tests are attractive, because such nulls may be obtained independent of any assumptions about tracing error other than that it be in-channel.

The test which is free of co-channel effects is that of pinch intermodulation. SMPTE tones are recorded in the lateral channel. Then the 1.1 sum and difference tones appearing in the vertical channel are characteristic only of in-channel perturbations. Unfortunately, the higher frequency carrier tone does not appear in the vertical channel so that a direct phase modulation test can not be made. That tone may be extracted from the lateral channel by means of a narrow pass-band filter for the purpose of a phase or frequency modulation test, but the filter must produce negligible phase shift, or the test would be vitiated. Otherwise, one makes the observation with the help of a wavemeter. As before, the sum of the sidetones, expressed as a percentage of the carrier strength, measures phase modulation only if that modulation is the dominant one. It is thus seen that the discrimination is only partial, so that mistaken interpretations are possible, especially if adjustments are made to seek a null so that amplitude perturbations may be left dominant.

Of the tests requiring a null for tracking error, those requiring a null for only the lateral error are likely to be the most convenient. The test for lateral error using rapid alternations of right and left crossmodulation tests so that the matching of these shows the lateral null, can be made to alternate with the tracing test. The lateral error is, of course, never large under normal circumstances, and the null can be found in advance to come at a particular groove radius, for a given adjustment of a pivoted tone arm.

With lateral tracking error nulled, one may look for FIM in the lateral channel, for laterally recorded tones in an SMPTE configuration. Wavemeter indication of the corresponding sidetones is an equivalent test. If an indication is obtained, the interpretation is that the stylus does not present the same magnitude of curvature to the two walls. Both this test and the pinch test require the sensitivity to the unwanted channel to be carefully cancelled out, using a mixing matrix if the cartridge provides insufficient separation.

There is an IM test for tracing, measuring the error for each of the right and left channels separately, which is insensitive to vertical tracking error. For this test, the low-frequency tone is recorded in the lateral channel but the high-frequency tone is recorded in one or the other of the right or left channels. The need to discriminate against amplitude intermodulation is met by using an SMPTE con-

figuration of test tones and observing the FIM signal in the channel in which the high frequency tone had been recorded. When the lateral error is null, the FIM measures only in-channel delay perturbations. Tracing error for each of the two channels may be separately measured and compared. This test would be especially interesting if the recording were made using RCA's Dynamic Recording Correlator, for then one could see if the observed insensitivity¹³ to playback stylus curvature would hold with a test capable of showing a sharper null with variations in curvature.

EXPERIMENTAL CHECK

As a check on the crossmodulation predictions of this paper, an experimental test record was made¹⁴ to provide an informal experimental verification. A more extensive report will be made at a later time.

The check is informal in another way, as there was no desire to disturb a productive studio facility for making the record. Thus the normal RIAA recording equalization was not disabled, nor was any special adjustment of the recording head undertaken beyond routine studio procedure. The head was a Westrex 3C cutter mounted with a CBS wedge and fitted with a CBS stylus, one which had already given some hours of service. Though these conditions made for some hardships in level calibration and controlling noise, they introduced an element of realism.

The measurements were made directly on the lacquer recording *via* playback with the Shure M44-1, M44-7, M44-5, and V-15 cartridges with stylus radii of 1.0, 0.7, 0.5, and 0.2 mil, respectively. Of these, the greatest number of readings was made with the M44-5, so that these data were chosen for display in Fig. 9. Level calibration, using a standard test record for comparison, showed that a 400 cps tone had been recorded at a peak velocity of 4.13 cm/sec in the right channel, and a 4000 cps tone at 1.17 cm/sec had been recorded in the left, at a radius of 5.6 inches.

Distortion measurements were made using a conventional IM meter and also recording the strengths of the 3600 and 4400 cps side tones relative to that of the 4000 cps tone using a wavemeter. The signal was the velocity waveform produced by the left channel. The lateral tracking error was "nulled" by holding the axis of the cartridge body normal to the radius from the center hole, since the experimental recording provided for no convenient test for that error. The vertical tracking error was varied by means of a pivoted turntable mount described elsewhere.¹⁵

In Fig. 9 the IM meter data are plotted, the symbol being an open diamond and the label being AXM (amplitude crossmodulation). The data show no particular dependence on inclination, but do indicate that the amplitude modulation is smaller than the sum of the sidetone strengths. This verifies the interpretation of the sidetone sum as indicating phase crossmodulation. Those sums are plotted with a solid circle as symbol and labeled TXM (total crossmodulation).

Readings were taken for the turntable tipped at angles of 0° , 2° , 4° , 6° , 8° , 10° , and 12° from its normal position, the positive sense being the one which increased the clearance

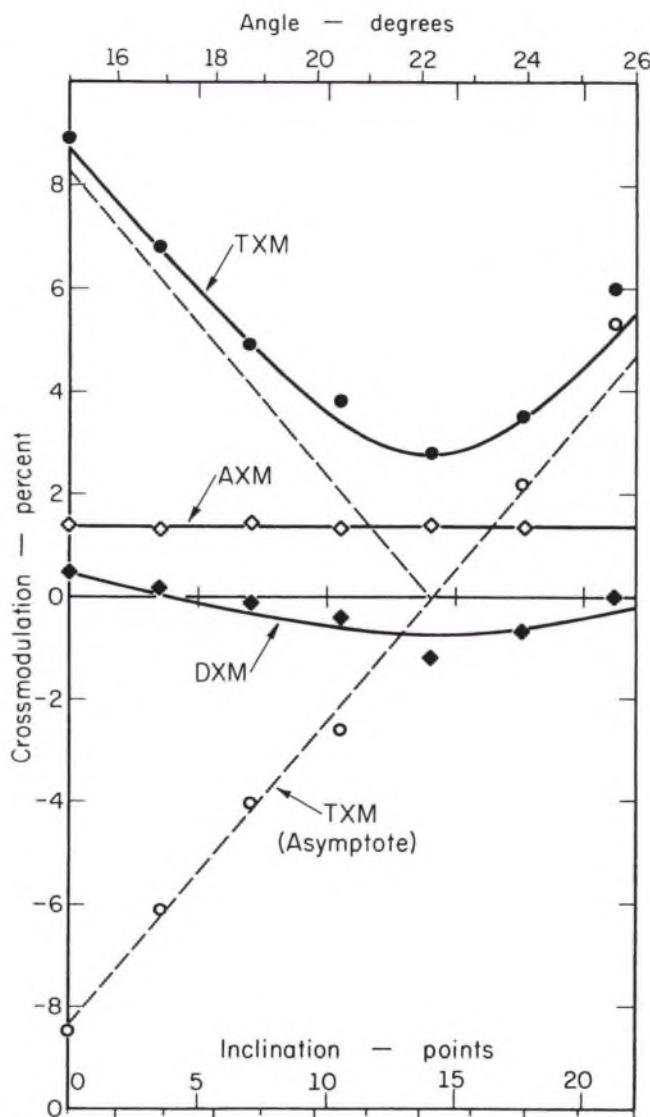


FIG. 9. Experimental crossmodulation check. The test recording placed a 4.13 cm/sec 400 cps tone in the right channel and a 1.17 cm/sec 4000 cps tone in the left, at a radius of 5.6 in. The solid circles mark the data taken as the sum of the 3600 and 4400 cps sidetones appearing in the velocity waveform from the left channel, expressed as a percentage of the 4000 cps carrier in that same channel; the sum represents total crossmodulation (TXM). The AXM data read from a conventional IM meter show the amplitude crossmodulation to be dominated by noise and phase crossmodulation. The same wavemeter data used for TXM is less affected by noise; the difference between the upper and lower sidetones should indicate an amplitude crossmodulation. The data are labelled DXM, and show more modulation than theory predicts. The open circles for the TXM data show the effect of removing the minimum-value quadrature contribution by root-difference-of-squares. The slope of the asymptote is determined from the recording level, and its zero intercept for a best fit to the data. The TXM curve is the theoretical hyperbola, for the point scale used, constructed for these asymptotes and minimum value. The inclination is that of the Shure M44-5 cartridge body. The angle scale is based on a 15° inherent angle for this cartridge. The cutter was a Westrex 3-C with the CBS modification. As mentioned in text, the data do not imply a calibration of this cutter.

¹⁴ The record was made in the studios of the Universal Recording Company of Chicago, Ill., under the sponsorship of Shure Brothers, Inc. Mr. C. R. Anderson of Shure Brothers supervised the making of the record and the subsequent measurements.

at the terminal end of the cartridge. The tangent of these angles is the horizontal scale, the inclination quoted in points. The minimum TXM is observed at 8° , an inclination of 14.05 points; the minimum value is 2.8 percent. The TXM curve is a theoretical hyperbola fitted to the observed data.

The fit was obtained by assuming the minimum value to represent phase crossmodulation in quadrature with that due to tracking error, so that it could be removed by a root-difference-of-squares technique. The results of this operation are plotted by open circles in Fig. 9, where an assignment of sign is also shown. The datum, so reduced, corresponding to the inclination of 14.05 points, is not plotted, since that datum is "used up" in reducing the others. The open circles should fit a straight line whose slope is already fixed, according to tracking error theory, by the recording level for the 400 cps tone. That straight line is the TXM asymptote shown dashed; its zero intercept was chosen, by eye, for the best fit. This line was the basis for the erection of the second asymptote, and, by a root-sum-of-squares calculation, the construction of the hyperbola. Since the distortion measurements may be trusted to about ± 0.2 percent of the carrier strength and the angles to about ± 0.2 degrees in any one sequence, the fit is very satisfactory.

The M44-5 is nominally a 15-degree cartridge. The angle scale shown at the top of Fig. 9 was computed with the assumption that zero inclination of the cartridge body corresponded to an inherent vertical tracking angle of 15° or 26.8 points of interent inclination. The minimum is then shown as indicating a cutting angle of 22.2 degrees, agreeing with another observation for a different specimen of the CBS modification.¹

The agreement is fortuitous and contains no implication for current practice at CBS. The nominal accuracy in the present observation represents a potential not fulfilled in this informal test because no basis was provided for a calibration of comparable accuracy. The M44 cartridges were designed as high-performance cartridges, conforming, as closely as the state of the art would allow, to the 15 degree standard, but they were not designed to serve as standard "protractors." The state of the art allowed a design accuracy of perhaps 3 degrees at best. Also, in the present experiment, the knowledge of the vertical error is contaminated by any residual lateral error. Thus, it was not surprising to find that the indications obtained from the M44-1, M44-7, M44-5, and V-15 disagreed with one another by some 2 to 4 degrees, though there was no disagreement between the M44-7 and the V-15.

Since the measurements were made in the velocity waveform, there should be observable amplitude crossmodulation due to tracking error. The IM meter readings give no such indication, the reason being that the experimental recording showed an excessive noise level compared to the one percent or less of expected AXM. The narrower bandwidth of the wavemeter places it in a more advantageous position with respect to noise. Miss-tuning the wavemeter allowed a rough estimate of the noise level as being 0.4 percent of the strength of the 4000 cps carrier.

For this reason, a more sensitive test for amplitude cross-modulation is the observation of the difference between upper and lower sidetones, assuming the upper sidetone to be the stronger, as theory predicts. These data are labelled DXM (difference crossmodulation) and are plotted as solid diamonds in Fig. 9. The DXM curve is the same hyperbola plotted above, but with the vertical scale compressed in the ratio 5:1. The theoretical compression should be 10:1. Thus, the observed amplitude crossmodulation is at least twice that to be expected from theory. Also, the data show the lower sidetone nearly always to be the stronger. It is not possible, at this writing, to explain these anomalies as

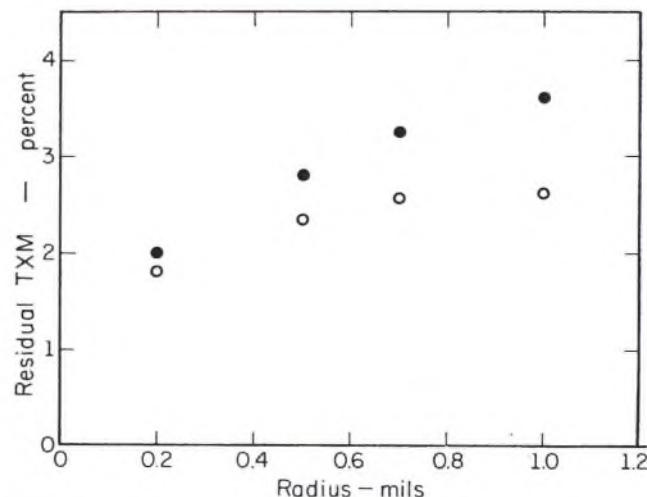


FIG. 10. Residual crossmodulation. The minimum values from data like those of Fig. 9, but taken using the Shure M44-1, M44-7, M44-5, and V-15, are plotted vs the stylus radius of curvature. Some part of the total residuum, plotted by solid circles is due to tracing distortion appearing because of cutter crosstalk, but only a small part, because this crosstalk was more than 20 db down. Assuming the 20 db figure, the amount to be expected was calculated and subtracted from the total, leaving a reduced residuum plotted as open circles. The reduced residuum is everywhere at least a factor of three greater than the estimated noise. There is not enough information in these preliminary data to allow the residuum to be identified. The total residuum blunts the null of Fig. 9 so that it does not fulfill the promise of Fig. 8, even taking the reduced recording level into account.

due to noise, to systematic instrumental errors, or to an unfortunate scatter in the data. Such explanations must await a more refined experiment.

Nor are the TXM data, though indicative of phase cross-modulation, free of anomaly. The data for the other cartridges, not shown, tend to require a more steeply sloping asymptote than justified by the recording velocity for a convincing fit, although the data are too few to make such a judgement definitive. The more serious anomaly is the excessive residual distortion at the minimum in the TXM curve. The amount is not consistent with the residual tracing distortion to be expected because of crosstalk between channels. The crosstalk was very small, down by at least 20 db.

The effect of the expected tracing distortion, based on the 20 db figure, may be seen in Fig. 10 where residual TXM

is plotted vs the playback stylus radius. The solid circles mark the total residuum, and the open circles show the result after subtracting the expected tracing distortion. It is seen that very little of the residual distortion may be attributed to tracing error. These data are not independent of the DXM data, so that it may be that the anomalies noted there are connected with the present ones. It is not possible to rule out some instrumental artifact, or even noise, as accounting for the residuum, though there is no convincing evidence for such explanations. Independent evidence, such as direct observation of frequency or phase crossmodulation, would be needed. (See *Letter to the Editor* in this issue of the *Journal*.) Should the residual distortion remain obdurately real, it could represent a largely co-channel delay-perturbing nonlinearity arising from an axial motion of either the cutting or reproducing stylus, or both, but in phase quadrature with the motions of tracking distortion.

In any case, something has blunted the tracking error null, so that the full theoretical acuity is not shown in the present experiment. The null is less than a factor of two sharper than would have been shown in an intermodulation test at the same levels. It is to be hoped that future experiments will either show sharper nulls or provide an identification of the residual distortion. Either result would signal an advance in the art. For the latter, the advance would be the availability of a more incisive tool for the study of a hitherto unobserved mode of distortion, leading to means for its control.

ACKNOWLEDGEMENTS

The author has had the benefit of many stimulating discussions with J. H. Kogen, C. R. Anderson and J. E. Jacobs of Shure Brothers, Inc. The opinions and proposals in this paper are, however, the author's own and are not necessarily those which would have been offered by Shure Brothers. The author would like to thank Mr. S. N. Shure for providing conditions favorable to these studies.

REFERENCES

1. J. B. Halter and J. G. Woodward, "Measurement of Distortions Due to Vertical Tracking Angle Errors in Stereodisk Systems," *J. Audio Eng. Soc.* 12, 8 (Jan., 1964); also, J. B. Halter, "Further Comments on the Measurement of Distortions in Stereodisk Systems," *J. Audio Eng. Soc.* 12, 162 (L) (Apr., 1964).
2. W. D. Lewis and F. V. Hunt, "A Theory of Tracing Distortion in Sound Reproduction from Phonograph Records," *J. Acoust. Soc. Am.* 12, 348 (Jan., 1964).
3. M. S. Corrington, "Tracing Distortion in Phonograph Records," *RCA Rev.* 10, 241 (1949).
4. H. Fletcher, *Speech and Hearing in Communication* (D. Van Nostrand Company, New York, 1953), pp. 153-175.
5. J. Dugundji, "Envelopes and Pre-Envelopes of Real Waveforms," *IRE Trans. on Info. Theory* IT-4, 53 (1958).
6. E. C. Fox and J. G. Woodward, "Tracing Distortion—Its Cause and Correction in Stereodisk Recording Systems," *J. Audio Eng. Soc.* 11, 294 (1963).
7. D. H. Cooper, "Unified Analysis of Tracing and Tracking Error," *IEEE Trans. on Audio AU-11*, 141 (1963).
8. D. H. Cooper, "Integrated Treatment of Tracing and Tracking Error," *J. Audio Eng. Soc.* 12, 2 (Jan., 1964).
9. B. B. Bauer and A. Schwartz, "A Record for Measuring Vertical Pickup Angles," Fifteenth Annual Convention of the Audio Eng. Soc. (October 17, 1963).
10. D. H. Cooper, "Tracking Distortion as Phase Modulation," *IEEE Trans. on Audio AU-11*, 41 (1963).
11. B. B. Bauer, "Vertical Tracking Improvements in Stereo Recording," *Audio* 47, 19 (Feb., 1963).
12. J. G. Woodward and E. C. Fox, "A Study of Tracking Angle Errors in Stereodisk Recording," *IEEE Trans. on Audio AU-11*, 56 (1963).
13. J. G. Woodward, "Reducing Distortion in Stereo Phonograph Systems," *Audio* 48, 23 (Feb., 1964).
14. R. W. Bayliff, "Some Aspects of Gramophone Pick-Up Design," *Proc. Inst. Electrical Eng. (B)* 109, 233 (1962).
15. J. Rabinow and E. Cordier, "Phonograph Needle Drag Distortion," *J. Acoust. Soc. Am.* 24, 216 (1952).
16. M. S. Corrington and T. Murakami, "Tracing Distortion in Stereophonic Disc Recording," *RCA Rev.* 19, 216 (1958).
17. J. Walton, "Stylus Mass and Reproduction Distortion," *J. Audio Eng. Soc.* 11, 104 (1963).
18. L. E. Goodman, "Contact Stress Analysis of Normally Loaded Rough Spheres," *J. Applied Mech. (Trans. of the ASME, Series E)* 29, 515 (1962).
19. J. Kogen and R. Samson, "The Elliptical Stylus," *Audio* 48, 33 (May, 1964).

LETTERS TO THE EDITOR

Note to Members: This is your column. It is designed for the discussion of papers published in the Journal and other pertinent topics about which you feel strongly.

EPILOGUE ON MEASUREMENTS

DUANE H. COOPER

University of Illinois, Urbana, Illinois

IN the time since the article *On Tracking and Tracing Error Measurements*¹ was written, I have had the good fortune to receive comments from several persons who had read advance copies of the manuscript. As is often the case, many of the comments were from referees, and contained valuable suggestions for improving the article. It was not possible to take all the suggestions into account in the time available for getting the manuscript to the printer, however, and I would like to amplify here those remaining points which have been called to my attention.

One of these points concerns the presentations in Figs. 8 and 9 of that article, in relation to the standardization of frequency-modulation data. It was observed for Fig. 8 that plotting one curve for standardized values, and the other not, could lead the reader to some unfair misinterpretations, and that the relative sharpness of the null in Fig. 9 was difficult to judge in the absence of a reference curve. Accordingly, I have prepared Fig. A as a supplement to both Fig. 8 and Fig. 9.

The conditions were set up as for the experimental test recording. For the FXM and TXM curves, a 400 cycle signal at 4.13 cm/sec is assumed recorded in one channel and a 4000 cps tone in the other, at a groove radius of 5.6 in. For the FIM curve, vertical recording, resulting from the two groove walls each being modulated at 2.06 cm/sec, is assumed for the 400 cycle signal. For the FXM curve, a 26 db cutter crosstalk figure is assumed; thus, a 400 cycle leakage signal at 0.206 cm/sec will appear in the 4000 cycle channel, giving rise to some tracing intermodulation, calculated for a 1.0 mil stylus to be the quadrature component at the minimum. For the FIM curve, the calculated quadrature component arises from tracing error for the same stylus, but for a 2.06 cm/sec 400 cycle signal. The TXM curve is that actually fitted to the experimental data for a 1.0 mil stylus.

The three curves all represent the same standardized manner of quoting a frequency-modulation index: frequency deviation as a percentage of the modulating frequency for FXM and FIM, or the exact equivalent for small modulation index, namely the sum of the first-order side tones for TXM. Also, it will be observed that the 400 cycle vertical

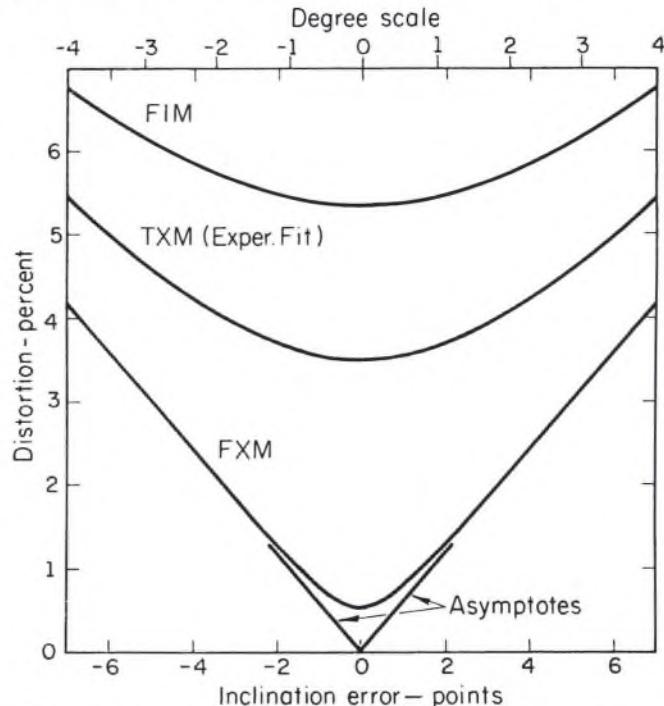


FIG. A. Comparison of theoretical intermodulation and cross-modulation measured by standard FM indices with experimental crossmodulation measured by sum-of-sidetones, as functions of tracking error inclination.

component of the groove-wall modulation has been taken to be the same in all three cases, 2.92 cm/sec. Thus the curves have identical tracking distortion asymptotes, and the sharpness and depth of the nulls may be fairly compared.

The claimed 10:1 ratio between the theoretical FIM and FXM null is satisfactorily displayed, and the blunting of the experimental TXM null is plainly evident, even though

¹ See p. 312 of this issue.

the selection of the 1.0 mil stylus places the TXM null in the more favorable light. The same blunting (or more) is seen in FXM data from this test recording, in observations made since the article was written, and this blunting has also been observed at RCA² in their FXM data, but showing a null some three to four times as deep as in FIM data. The cause of this residual distortion is as yet unknown, but the conjecture that it arises from a quadrature component in the springback motion of the cutter is quite tempting.

Quadrature springback distortion is distinguishable from tracing distortion in being a strictly co-channel phenomenon. Also, it should vary as the inverse first power of the cutting radius rather than as the inverse second power. Thus, it should make a weaker contribution to the quadrature component for a crossmodulation test at the shorter radii, and the null should be deeper, relative to the FIM null, at those radii, if that be the blunting factor. The present test, at a radius of 5.6 in., happens to give quadrature springback its best chance to put in an appearance, relative to tracing distortion. Again, quadrature springback should be sensitive to the stiffness of the cutting stylus in the direction of the groove axis. The modifications recently reported for the Westrex 3C cutter, resulting in the 3D version³ should show an enhanced stiffness leading to a substantial reduction in quadrature springback. Thus, if the present conjecture should prove valid, the crossmodulation test would be especially valuable for that cutter or any others showing an enhanced stiffness.

Another point on which I have been queried is my strong espousal of a standard way for reporting frequency modulation indices. It is perhaps unfortunate that I have chosen the word "standard" to describe what is to my mind more a matter of correctness and meaningful reporting than a matter for action by some standards committee. No committee can legislate correctness any more than it could legislate (to take a pretentious example) the value of π . In the intermodulation or crossmodulation experiment, there are two experimental frequencies with which the frequency deviation may be compared. Expressing the deviation as a percentage of the higher (carrier) frequency is one comparison, but the question is whether that is the more meaningful comparison of the two that may be made.

It turns out, upon Fourier analysis, that the relative intensity of the sidetones depends not only on the amount of the deviation, but also on the rate at which the deviation is undertaken, and that it is just the ratio of the deviation to the frequency at which the deviation occurs that counts. In fact, for small ratios, the relative side-tone strengths, summed together, are equal to that ratio. This is one of the reasons why it has not been customary, in FM communications, to compare the deviation with the carrier frequency. To quote an index for an 88 mc channel different from that for a 108 mc channel, when both use the same

deviations, would not be meaningful. There, the custom is either to use the true index, as given meaning in Fourier analysis, or to quote deviations in comparison with a maximally allowed deviation.

In psychoacoustics, the practice has grown of expressing a frequency deviation in comparison with the "carrier" frequency in dealing with the problem of flutter. I do not know why this is so, unless the practice was guided by the thought that the pitch discerning properties of the ear were being challenged. Certainly, that is more nearly the challenge for slow deviations, but it has been shown that for rapid deviations, the audibility of flutter may not be predicted without taking into account the rate at which the frequency deviations are undertaken.⁴ Indeed, the evidence strongly suggests that the application of masking theory in predicting the audibility of the FM sidetones in the presence of the carrier would account for the audibility of rapid flutter. In this way, the psychoacoustic significance of rapid FM would be properly suggested by the index measuring sidetone strength to the same degree that the usual AM index suggests psychoacoustic significance. That index for FM is the one comparing the frequency deviation to the frequency with which the deviations are undertaken.

As it happens, for both tracing and tracking distortion, the deviation expressed as a percentage of the carrier frequency, the flutter index, represents the automodulation of the lower-frequency tone, as determined with the help of a higher-frequency pilot tone. The true crossmodulation or intermodulation index is greater than the flutter index by the ratio of the frequencies used. As pointed out already in 1941 by H. G. Baerwald,⁵ the situation is remarkably different from the "usual" amplitude perturbing nonlinearities. Instead of roughly comparable intermodulation measure and harmonic content, the former is larger by an order of magnitude or more, depending on the ratio of frequencies. The warning is clear that the use of harmonic content or a flutter index seriously underrates the severity of the nonlinearity. Though the correct explanation appeared so long ago, one may still read that tracking distortion "is mainly second harmonic."

The quotation of a flutter index may not be quite so severely criticized today, provided it is cited for what it is, rather than being offered as a true FM index. The reason is that the reader will probably have been sensitized to the extremely low perception thresholds that have been shown for that phenomenon. Even so, the practice is to be discouraged; the reader may yet think that it is pitch discernment that is being challenged, and be led into confusing arguments such as those that once raged about Doppler distortion in loudspeakers. The specialists will not be confused, of course, but the practice would be a disservice to the able audio engineer whose responsibilities are so far reaching as to allow him to be a specialist in only a few topics.

² J. G. Woodward, private communication.
³ C. S. Nelson and J. W. Stafford, *J. Audio Eng. Soc.* 12, 178 (July, 1964).

⁴ F. A. Comerci, *J. Soc. Motion Picture Telev. Engrs.* 64, 117 (1955).
⁵ H. G. Baerwald, *J. Soc. Motion Picture Telev. Engrs.* 37, 591 (1941).

A New Method of Disc Recording for Reproduction with Reduced Distortion: the Tracing Simulator*

HORST REDLICH AND HANS-JOACHIM KLEMP

Telefunken-Decca, Berlin, Germany

By means of an analog-computer technique, geometric playback distortions are simulated during disc recording. Second- and third-order harmonics are recorded so that the vertical and lateral components are free of tracing distortion in playback and pinch-effect distortion is removed in lateral recording. The method also offers the possibility of correcting the vertical recording angle by electrical means.

INTRODUCTION

WHEN stereophonic records were introduced in 1958, the majority of record players were sold with stereo pickups, so that today both stereo and mono records are for the most part replayed with this kind of equipment. Since a stereo cartridge is an electromechanical two-component transducer, as a basic principle the geometrical sum of the output voltages of the two channels must be constant for a given displacement of the replay needle tip, no matter what the direction of displacement. This means that all distortions caused by the geometrical relationship between the groove, its displacements and the sapphire replay needle tip are fully effective on reproduction, including those components which are active vertically in relation to the displacement. A single-component pickup used purely for lateral recordings is insensitive to a large proportion of the geometrical replay distortions.

The record industry has recently been concerned with overcoming these defects in reproduction with a stereo pickup by using special recording methods. The method of recording described in this article eliminates practically all additional distortions caused by the use of two-component cartridges. This applies both to stereo and to mono recordings. This method offers the possibility of compensating for square and cubic distortion components. Consequently the level on records cut using this method can be increased, or, for a given recording level, the groove speed reduced.

What exactly are geometrical replay distortions? In the years 1937-1949 studies¹⁻⁴ were already published investigating those distortions that result from the fact that the center of the replay half-ball does not precisely follow the groove displacements; these are referred to as tracing distortions and distortions resulting from pinch effect. Furthermore, geometrical replay distortions may be due to the fact that the plane of movement of the recording system

does not coincide with that of the replay cartridge.^{5,6} The angle between the two planes is known as the tracking angle error.

TRACING DISTORTIONS AND DISTORTIONS CAUSED BY PINCH EFFECT

Figure 1 shows a vertically recorded sinewave and replay

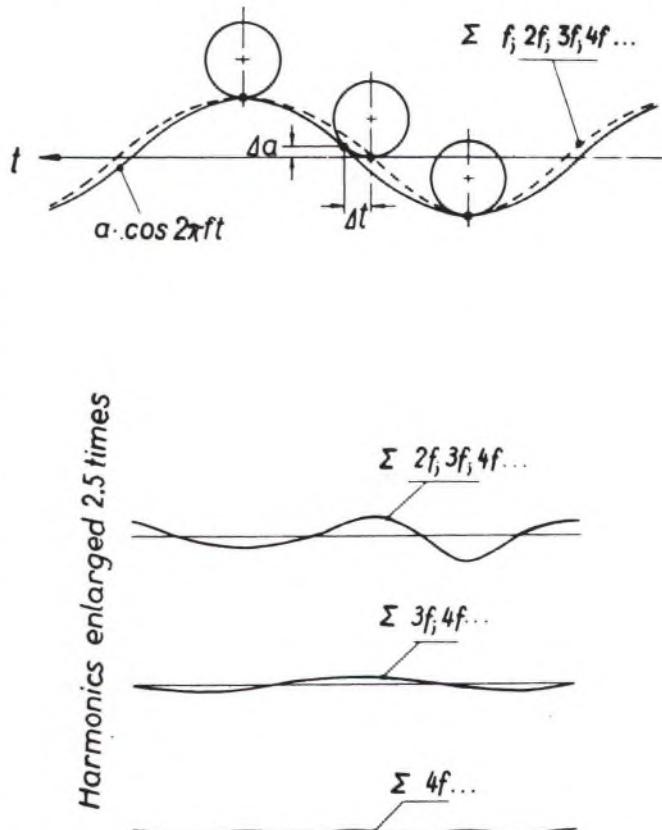


FIG. 1. Tracing distortions for the replay of vertical recording.

* Presented October 13, 1964 at the Sixteenth Annual Fall Convention of the Audio Engineering Society, New York.

curve which deviates from the recording. In the case of lateral replay, Fig. 2, it is clear that if the centerline of the recording is a sinewave, the lines of contact of the replay half-ball with the two groove walls again correspond to the distorted replay curve, except that the upper and lower curves are shifted through 180° in relation to each

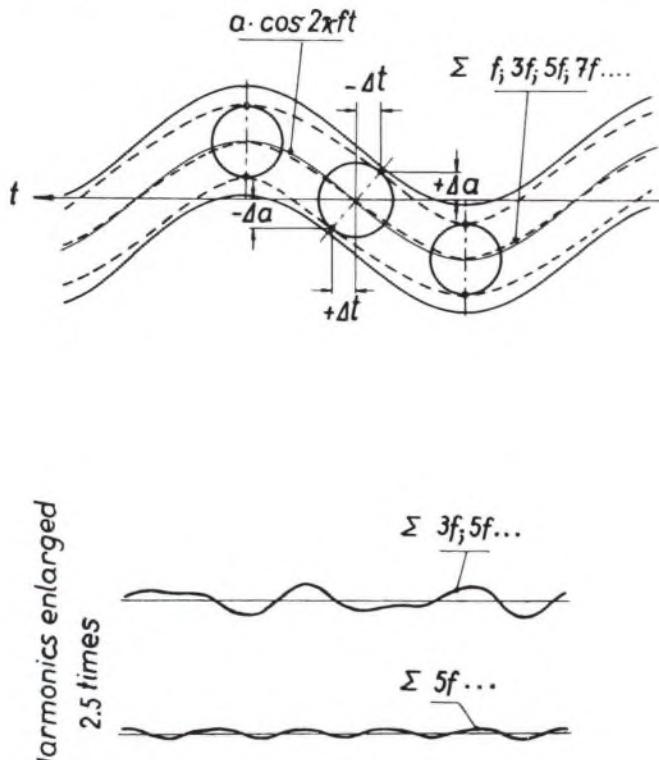


FIG. 2. Tracing distortion for the replay of lateral recording.

other. The result is that where the signal line passes through zero, the Δa value on the one hand and the Δt value on the other cancel out, and consequently the even order components of the tracing distortions are eliminated in the case of lateral recording.² Lateral recording is therefore also known as push-pull recording. The pinch effect distortions are shown in Fig. 3. In Table I the harmonics are arranged in accordance with the recording components and the direction in which they act as distortion during replay.

Quantitative evaluation shows that with higher order of harmonics the distortions diminish rapidly. Generally the second harmonic causes interference, but under certain conditions the third harmonic must also be taken into account. According to Corrington,⁴ the amplitudes of the tracing distortions in the case of velocity-dependent pickup are calculated as follows:

Second harmonics in the case of vertical recording and

second harmonics in the case of lateral recording caused by pinch effect

$$= (1/2)RA^2[1-(1/4)A^2+(15/128)A^4-\dots]- \\ (1/6)R^3A^4[1-(15/16)A^2+\dots]+ \\ (1/48)R^5A^6(1-\dots)-\dots$$

Third harmonic in the case of vertical recording

$$= (3/8)R^2A^3[1-(3/4)A^2+(9/16)A^4-\dots]+ \\ (27/128)R^4A^5[1-(3/2)A^2+\dots]- \\ (243/5120)R^6A^7[1-\dots]+\dots$$

Third harmonic in the case of lateral recording

$$= -(3/16)R^2A^3[1-(3/8)A^2+(9/64)A^4-\dots]+ \\ (27/512)R^4A^5[1-(3/4)A^2+\dots]- \\ (243/40960)R^6A^7[1-\dots]+\dots$$

where $A = (2\pi \cdot a)/\lambda$, $R = (2\pi \cdot r)/\lambda$, and a = recording amplitude, λ = wavelength, r = radius of replay needle tip.

In the case of a long-playing record these complicated

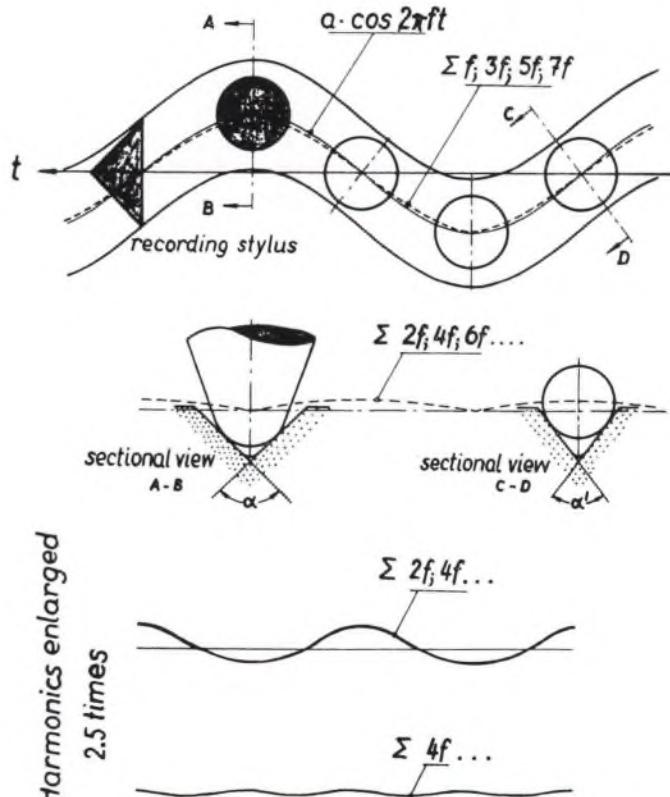


FIG. 3. Pinch effect distortions for the replay of lateral recording.

expressions can be simplified considerably, if an error of ≤ 1 db is admissible in the result.³ The first member of the series is then sufficient; consequently, for the recording

TABLE I. Harmonies of geometric replay distortions.

Recorded signal	Description of distortions	Order of harmonics	Direction of distortions during replay
Vertical	Tracing distortions	$2f, 3f, 4f, 5f, 6f \dots$	Vertical components
Lateral	Tracing distortions	$3f, 5f \dots$	Lateral components
Lateral	Pinch effect distortions	$2f, 4f, 6f \dots$	Vertical components

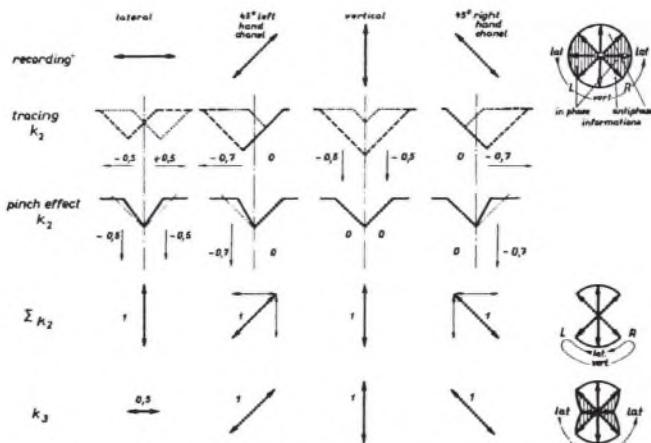


FIG. 4. Direction of geometric replay distortions.

of a sinewave signal $a \cdot \sin \omega t$, the velocity of the second harmonic in the case of vertical recording and caused by pinch effect

$$= \pi \cdot r \cdot f \cdot (s^2/v^2) \sin 2\omega t. \quad (1)$$

Third harmonic in the case of vertical recording

$$= (2/3) \pi^2 \cdot r^2 \cdot f^2 \cdot (s^3/v^4) \cos 3\omega t. \quad (2)$$

Third harmonic in the case of lateral recording

$$= (3/4) \pi^2 \cdot r^2 \cdot f^2 \cdot (s^3/v^4) \cos 3\omega t, \quad (3)$$

where s = velocity of fundamental, v = groove speed, f = recording frequency.

The square components of the tracing distortions increase in linear proportion with the radius of the replay tip and the frequency, but the increase is proportional to the square of the increase in level and reduction in groove speed. The cubic components, on the other hand, increase with the square of the replay tip radius and the frequency, whereas they increase with the cube of the level and decrease with the 4th power of the groove speed. According to Lewis and Hunt³ this observation also applies to intermodulation and two-tone distortions.

DIRECTION OF DISTORTIONS

When replaying stereosignals in two-component recordings the plane of movement of the replay needle tip forms a cycle depending on the phases of the two informations to each other, assuming that the two amplitudes are constant. But this is not valid for the harmonics. Figure 4 shows the vectors of distortions where the unit of amplitude was selected as one. The second harmonic is the geometric sum of tracing distortion and distortions caused by the pinch effect, to which the distortions of the two groove walls should also be added. Therefore, in the quadrants of in-phase stereo informations, no second-order harmonics are generated. The plane of movement of this kind of distortion concerns only the range of anti-phase information, with a limit of one channel modulated only. The third-order harmonic is given by the tracing distortions only, and their direction corresponds in every case to that of the fundamental. But in the direction of the lateral component the

amplitude of the third harmonic is only half of that for the vertical component (see Eqs. 2 and 3). Therefore, the amplitude of third-order distortions decreases between 45° and lateral recording, with $a_3 = 1/2(1+\tan a)$, a = angle of the recording, while it is constant for recording at the angle between single-channel and antiphase modulation.

DISTORTION DUE TO VERTICAL TRACKING ANGLE ERROR

Since the point of suspension of a system must always lie above the surface of the record, the movement of the pickup shown in Fig. 5 as a tangent on the arc is not directed vertically to the surface of the record but inclined by the angle α in the direction of rotation of the groove. The result is that the center of the replay half-ball again does not follow the groove displacement if the recording is cut at another angle, e.g., vertically. Distortions due to the vertical tracking angle error δ , i.e., to the difference in angle between the tracking angle of the recording γ and the tracking angle of the replay cartridge α , take the form of mainly even-order harmonics. The phase position of the second harmonic relative to the fundamental is displaced through 90° compared to the second harmonic of tracing distortions and the pinch effect. The velocity of the distortion caused by the tracking angle error for the second harmonic amounts to

$$\delta(s^2/v) \cos 2\omega t,$$

where δ is expressed in radians. Consequently, this distortion is directly proportional to the tracking angle error and the groove speed and to the square of the level.

The vertical tracking angle error should be taken in account not only for the replay system but also during the recording. By the springback of the lacquer material, during the cutting process the recording is deformed. B. B. Bauer⁶ shows that this deformation is equivalent to the distortions given by vertical tracking angle error. The

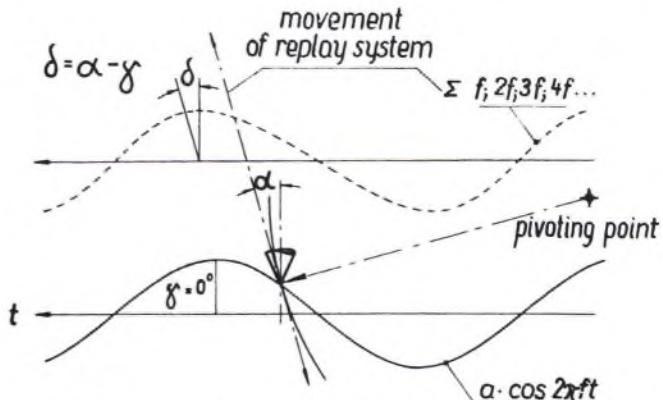


FIG. 5. Distortion of the movement of the replay tip due to the vertical tracking angle error.

springback of the acetate impairs the recording for vertical and for lateral modulation (as discussed under *Vertical Tracking Angle of the Recording*).

In conclusion, we can summarize the above remarks as follows: when a record is replayed with a stereo cartridge, additional, mainly second and cubic, harmonic distortions

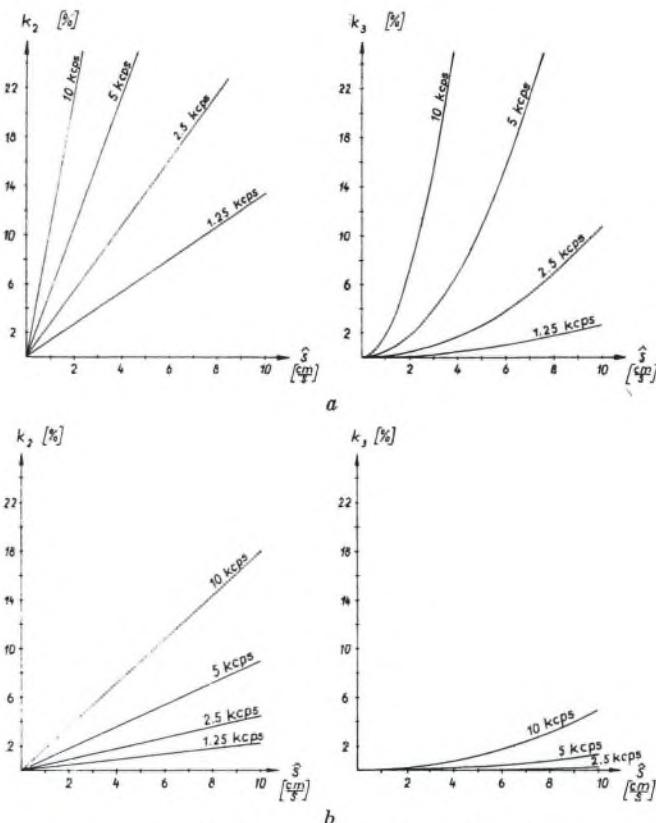


FIG. 6. Second and third harmonics for the replay of an LP record.
a. $n = 33\frac{1}{3}$ rpm, replay tip radius $r = 15 \mu\text{m}$, groove diameter $d = 290 \text{ mm}$; b. $n = 33\frac{1}{3}$ rpm, replay tip radius $r = 15 \mu\text{m}$, groove diameter $d = 120 \text{ mm}$.

occur which have different causes. Since the amplitudes of the individual components are comparable, significant improvement in the quality of reproduction is only possible if all three types of distortions are taken into account with regard to their magnitude and the direction in which they act during replay. Figures 6 and 7 give a survey of the geometric replay distortions for a $33\frac{1}{3}$ rpm LP record when replayed with a replay needle tip of $15 \mu\text{m}$ radius.

POSSIBLE METHODS OF COMPENSATING FOR DISTORTIONS

Various suggestions have been made for compensating the

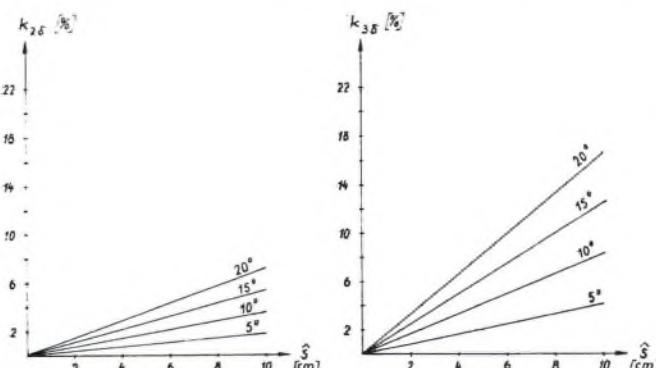


FIG. 7. Second harmonics due to the tracking angle error for $d = 290 \text{ mm}$ and 120 mm groove diameter.

geometrical distortions that occur during replay by means of a suitable predistortion of the recording. McNair observed in 1938 already that if the vertical recording were subjected to intermediate replay with a cartridge and then re-recorded again with reversed phase position, the tracing distortions with the same pickup would be reduced during subsequent replay. Fox and Woodward⁷ describe a method, used in practice for the first time by RCA, which compensates the tracing distortions in a stereo recording. The recording signals are phase-modulated in a Dynamic Recording Correlator by means of a delay line the time of which changes depending on the instantaneous value of the recording amplitude.

THE TRACING SIMULATOR

The purpose of the new method is to correct all the square and cube components of the geometrical replay distortions, taking into account the direction (i.e., vertical or lateral) in which they are replayed by means of appropriate predistortion during recording. This group includes tracing distortions, pinch effect, vertical tracking error, and spring-back of the two groove walls caused by the lacquer material.

It should be realized that the distorted curve shown in Fig. 1 can be changed into the undistorted curve not only by correction of the instantaneous values in the sense of the time axis, but also by altering the instantaneous values of amplitude. This is achieved by generating harmonics of

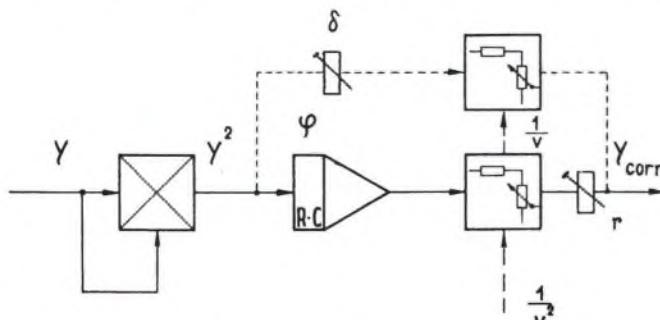


FIG. 8. Circuit diagram, function-former y^2 .

the signal to be recorded and adding them to the signal, in suitable amplitude and phase relations, to produce compensation. Since for this method the harmonics must be derived according to the laws which apply to the individual causes of distortion, they are available as a correction signal separate from the main signal. In this way it is possible to record them in every direction, i.e., the distortions associated with vertical recording in the direction of the vertical components, and in the case of distortions caused by lateral recording the even-order distortions (pinch effect) as vertical components and the non-even order distortions as lateral components (see Fig. 4). This method was developed in the Teldes laboratory and led to the Tracing Simulator.

The mathematical derivation of the harmonics is achieved here on the principle of an analog computer. Different function-formers are available for the square and cube distortion components. Figure 8 (solid lines) shows the circuit diagram of a function-former which calculates or derives,

from the input signal, the correction signal to compensate for the square component of the geometrical replay distortions for every instantaneous value of the recording amplitude. The first circuit arrangement causes the amplitude of the output signal to vary with the square of the input signal while at the same time doubling the frequency. In the following stage the signal is differentiated and phase-rotated through 90° while the linear increase with the frequency is also produced. The subsequent attenuator which alters attenuation with the groove diameter of the recording gives the square law relationship of the output voltage on the groove speed. Finally, the preset potentiometer acts as a constant factor which takes into account the radius of the replay half-ball. Thus analogy with Eq. 1 is achieved. Figure 9 shows the function-former used to derive the cube distortion components. Using this arrangement, an analogy is provided with Eqs. 2 and 3. While the square distortions can be derived directly from a characteristic which takes the shape of a square-function parabola, the cube components can only be obtained in a roundabout manner. The reason for this is that a sinewave $a \sin \omega t$ reflected in a square-function parabola gives the required expression

$$(1/2)a^2 \sin 2\omega t.$$

A characteristic which takes the shape of a cubic-function parabola, on the other hand, converts the input signal into

$$a^3 \sin^3 \omega t = (3/4)a^3 \sin \omega t - (1/4)a^3 \sin 3\omega t,$$

or, expressed differently: at the output from a circuit arrangement of this type 75% fundamental and 25% third harmonic will be available. To obtain the harmonic separately, $a^3 \sin \omega t$ must be eliminated. This is due to a parallel circuit for the cube characteristic. After twice differentiating the signal and passing through the output stage which gives the dependence on the groove speed, the correction signal to compensate for the cube distortion components is obtained. The Tracing Simulator contains two function-formers each of types y^2 and y^3 . When passed through suitable matrixing networks, the derived compensation signals are then superimposed on the recording signals in such a way that distortions that arise during replay and the distortions recorded by means of the Tracing Simulator eliminate each other. The direction of the respective distortions is taken into account by matrixing.

THE VERTICAL TRACKING ANGLE OF THE RECORDING

In order to ensure that the recording is made at the internationally agreed angle of 15° , the plane of movement of the recording system must be inclined through an additional angle to compensate for the springback effect, which as mentioned above depends on the elastic properties of the lacquer disc coating. The newly designed stereo cutter SX45/15s allows tilting up to 30° . The vibratory system is very rigid in the direction of movement of the sound groove, so that the adjusted vertical tracking angle does not depend to any great extent on the frequency. This system has also been developed with reference to the frequency band for recordings with the Tracing Simulator.

From the evaluation of recordings made with and without

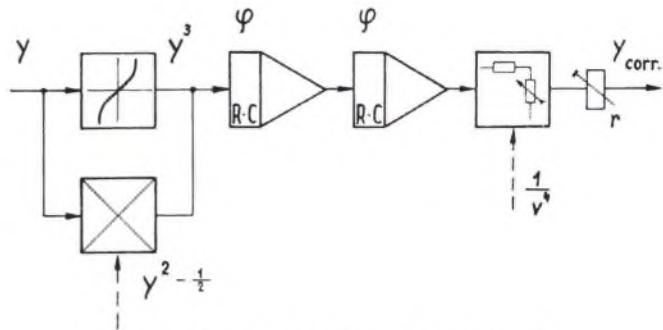


FIG. 9. Circuit diagram, function-former y^3 .

the simulator it results that the right and left groove wall have different springback effects. The direction of the replay distortions is then no longer coincident with the theory, as shown in Fig. 6. To obtain optimal compensation of replay distortions, the simulator is made in such a way that this recording error can be corrected. Therefore, an analogy to Eq. 4 is derived.

The signal generated by the frequency doubler in the function-former y^2 bypasses the differentiation stage fed into a separate attenuator which provides the necessary linear dependence on the groove speed. With the control δ the required correction can be adjusted (Fig. 8, dotted lines). The square components of tracing distortions and those of the tracking angle error are displaced at an angle of 90° to each other. This is taken into account by the geometrical addition of the two compensation signals. The correction signal to compensate for both distortions is then

$$U_{Corr} = \sqrt{(U_{Tr}^2 + U_\delta^2)}.$$

THEORETICAL AND ACTUAL DISTORTION COMPENSATION

The geometrical replay distortions of a longplay record are summarized in Figs. 6 and 7. The exceptionally high distortion for electroacoustical transmission with small groove diameters and high frequencies can be seen clearly.

Since, as described above, the replay distortions are simulated for correction, the degrees of compensation depends only on the accuracy with which amplitude and phase are reproduced. Figure 10 shows the error which arises in the simulator due to the simplifications made in the equations given above. The amplitude error for the square distortions is ≤ 0.5 db, and for the cube distortions is ≤ 1 db, i.e., the theoretical limit of compensation is ≥ 26 or ≥ 20 db.

This consideration is, however, only valid if the upper frequency limit of the recording is three times as high as that utilized during reproduction, and if within this range in the recording no phase shift takes place between the fundamental and the associated harmonics. Since a record cutter is not subject to phase shift within the negative feedback frequency range, but the phase of the recording changes outside the range of the negative feedback, the harmonic has a time lag when it is recorded. An important advantage is that in the case of the Tracing Simulator it is possible to retain the phase position by separating the fundamental and the harmonics with appropriate time delay simulation.

The maximum phase difference between the recorded fundamental and the harmonic can fairly easily be held

compensation signal [-dB]

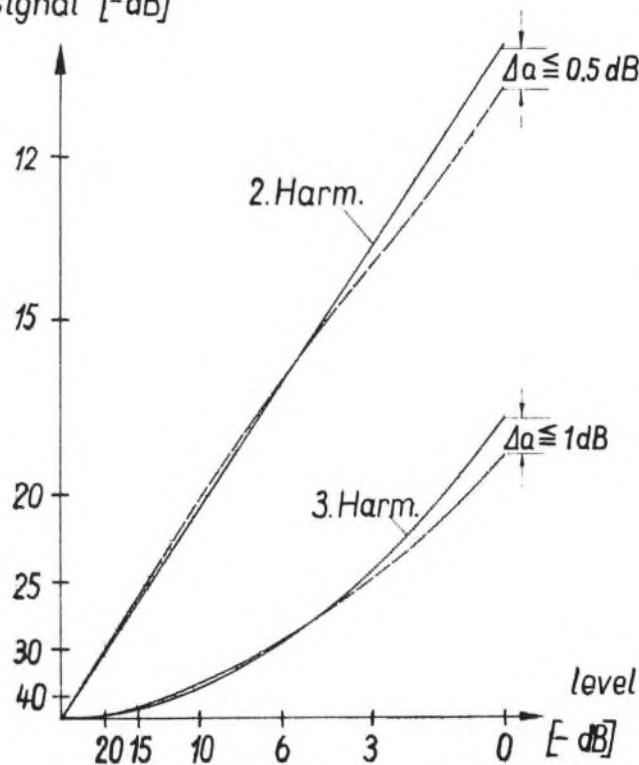


FIG. 10. Amplitude error of compensation voltages due to the simplified calculation in the simulator.

to $< 10^\circ$ over the whole frequency range used during replay. As a result, the minimum compensation of 15 db is obtained.

By correction of the first two harmonics which cause most interference, the straight section of the distortion vs level characteristic is extended almost to the theoretical limit. This limit is reached when the radius of curvature of the recorded signal is equal to the replay half-ball. The limiting level is calculated by the formula

$$S_{lim} = v^2/r \cdot w.$$

Fig. 11 shows the increase in the distortions below the

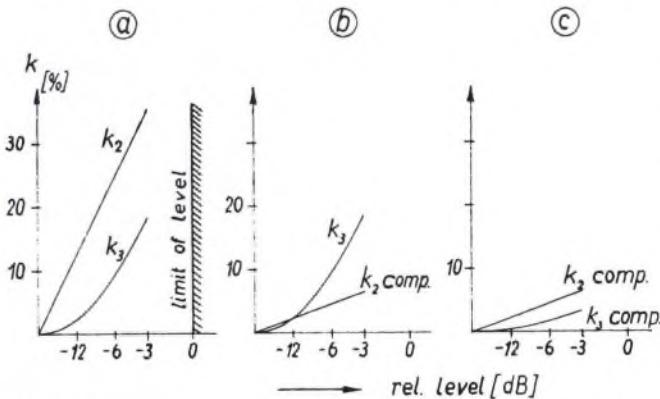


FIG. 11. Increase of distortion depending on the relative level. a. Conventional recording; b. only k_2 compensated (degree of compensation 15 db); c. k_2 and k_3 compensated (degree of compensation 15 db).

limiting level. If only k_2 is compensated, the curves of the percentage of harmonic distortion intersect at approximately -12 db, but they do not intersect if k_2 and k_3 are taken into account. In this case the point of intersection lies beyond the limiting level, in a range which for reasons of replay geometry cannot be utilized. By compensation of second and third order harmonics, the possible level is as near as 3 db to the theoretical level limit. In this range good reproduction quality is obtained, for practical purposes. Above this limit distortions increase rapidly, so that this range of recording level cannot be utilized. Thus, a significant improvement cannot be expected by the compensation of harmonics of still higher order.

Practical measurements coincide to a great extent with the theoretical results. Figure 12 compares the distortion

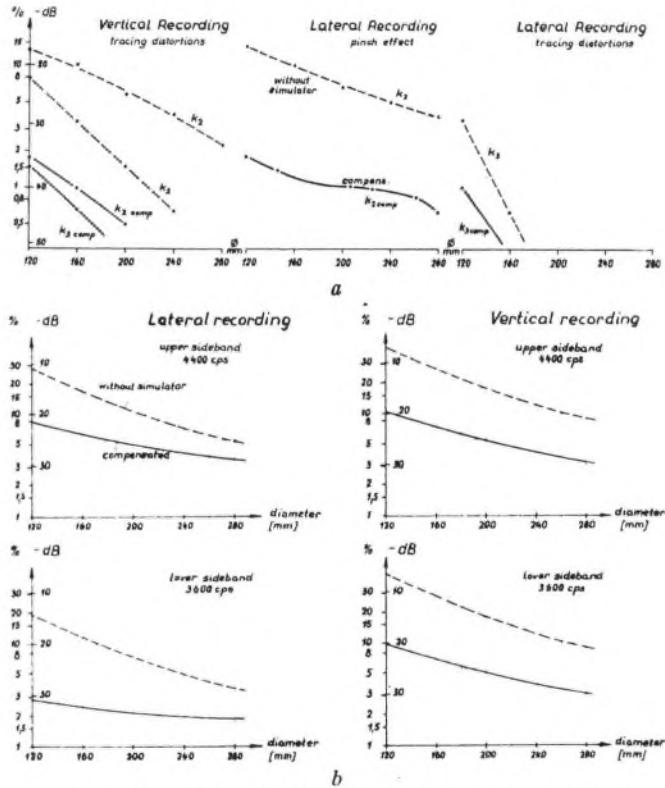


FIG. 12. Measured geometrical replay distortions of recordings made with and without the simulator depending on the groove diameter. a. One tone, $\dot{s} = 10$ cm/sec, $f = 3250$ cps. b. Intermodulation $f_1 = 400$ cps, $\dot{s} = 7$ cm/sec; $f_2 = 4000$ cps, $\dot{s} = 1.75$ cm/sec.

values measured during replay of an uncompensated recording with those produced by the Tracing Simulator. The amplitudes of the correction signals correspond to the calculated values, taking into account the linear replay losses. The selected level in the case of the smallest groove diameter lies only 3 db below the theoretical limiting level.

INFLUENCE OF THE VERTICAL TRACKING ANGLE ERROR OF THE PICKUP

For the measurements a pickup with 15° vertical tracking angle was used, corresponding to the recording angle. As mentioned above, the recording and replay planes must

coincide since a tracking angle error also leads to square distortions which considerably restrict the possible reduction in the tracing distortions. The practical significance of the relationship between tracing and angle distortions, the theory of which was established by Cooper,⁸ is shown in Fig. 13. So far the influence of the angle error has been small, due to the large tracing distortions, and only ascertainable at low frequencies; it hardly impaired the quality of reproduction.

When recording quality is improved using our method, however, it is of decisive importance. The deviation from

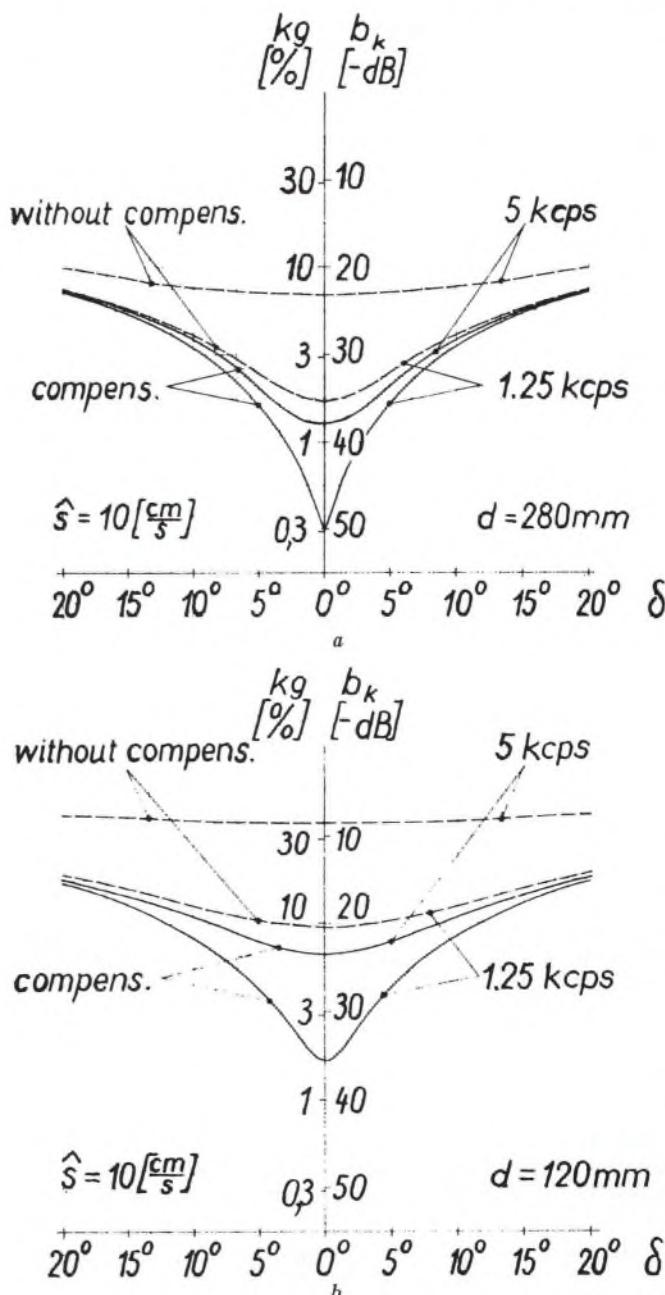


FIG. 13. Influence of the tracking angle error δ on the geometrical replay distortions in the case of recordings with and without compensation. a. Degree of compensation 15 db, groove diameter $d = 280$ mm. b. Degree of compensation 15 db, groove diameter $d = 120$ mm.

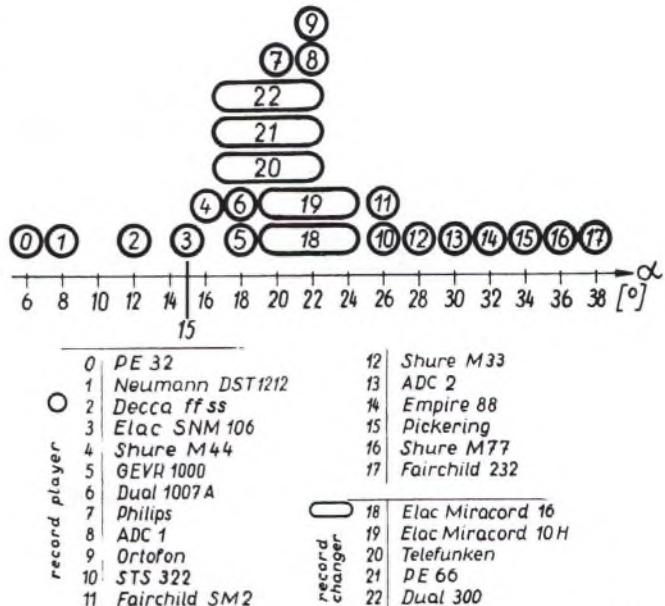


FIG. 14. Vertical tracking angle of different pickups in record players or changers.

the vertical tracking angle of 15° recommended as standard should not exceed 5° if good reproduction quality is desired.

Many pickups, especially in so-called Hi-Fi systems, are still far from meeting this requirement. Figure 14 gives the tracking angles for a large number of pickups measured in accordance with the provisions of the German Standards Institution, using CBS Test Record STR 160. The manufacturers of phonographs have for a long time been making efforts to maintain low horizontal tracking angle error for lateral recording, even in the case of the cheapest phonographs, by angling the pickup arm and placing it at a suitable position in relation to the turntable. It is now important to avoid *vertical* tracking angle error as well in the case of stereo recording.

INFLUENCE OF THE TOLERANCE OF THE REPLAY RADII

The dimension of the replay needle tip influences the value of replay distortion in accordance with the corresponding equations. Theoretically, small points reduce distortions. But there is a limit in practical use, since the specific needle pressure leads to deformation of the groove wall and thus again to distortions.⁹ Furthermore, it is impossible to produce grooves with extremely small-bottom radius. For the reproduction of recordings made by means of the Tracing Simulator the standardized replay needle tip-radius is the most suitable, as it gives the lowest distortion figure. As the radii of the replay tips vary between 12 and $18 \mu\text{m}$, the question arises as to what is the influence of a deviation from the mean value of $15 \mu\text{m}$. The calculation for the limit values, for example, gives a reduction in the degree of compensation from 15 to 12 db. The elliptical stylus being advertised at this time did not show essential improvements against the round ones in many measurements of normal and compensated recordings. Presumably the theoretically expected improvement is counteracted by the influence of higher specific loading of the groove walls.

The summary of all measurements is that recordings with compensation always result in smaller distortions if the tip radius is within the tolerances recommended by the standards. The amplitude of the compensation signal for recording should be referred to the mean value of the tip radii.

SUBJECTIVE IMPROVEMENT OF QUALITY BY THE TRACING SIMULATOR

To allow a subjective appraisal of the reduction in replay distortions, a number of different musical programs were recorded with and without the Tracing Simulator. Direct comparison of the reproductions clearly showed an improvement in quality in the case of small groove diameters, which could be noticed even by inexpert listeners. As was to be expected, the advantage was especially noticeable in the case of sound images which tend to form combination tones, such as choral singing, individual voices, or large string groups. For example, a recording of a choir of children had to be recorded at an 8 db lower level using the conventional method to give the same impression to listeners, as far as distortions were concerned. The general impression, however, was still not as good due to the lower signal-to-noise ratio. To check whether it would be possible to make full use of the improvement of quality with today's

pickup systems and to facilitate their further development, several test records will be made with the Tracing Simulator in conjunction with the German Standard Committee (DIN), and these test records will be checked in the expectation that it will be possible to develop a uniform method for distortion tests of pickup systems.

BIBLIOGRAPHY

1. M. J. DiToro, "Distortion in Reproduction of Hill and Dale Recording," *J. Soc. Motion Picture Engrs.* **29**, 493 (1937).
2. J. A. Pierce and F. V. Hunt, "Distortion in Sound Reproduction from Phonograph Records," *J. Acoust. Soc. Am.* **10**, 14 (1938).
3. W. D. Lewis and F. V. Hunt, "Theory of Tracing Distortion in Sound Reproduction from Phonograph Records," *J. Acoust. Soc. Am.* **12**, 348 (1941).
4. M. S. Corrington, "Tracing Distortion in Phonograph Records," *RCA Review* **10**, 241 (June, 1949).
5. B. Olney, "Distortion and Record Wear," *Electronics* **19**, (Nov., 1937).
6. B. B. Bauer, "The Vertical Tracking Angle Problems in Stereo Record Reproduction," *Electronics Industry Association* (February, 1963).
7. Fox and Woodward, "Tracking Distortion—Its Cause and Correction in Stereodisk Recording System," *J. Audio Eng. Soc.* **11**, 294 (1963).
8. D. H. Cooper, "Integrated Treatment of Tracing and Tracking Error," *J. Audio Eng. Soc.* **12**, 1 (1964).
9. F. V. Hunt, "Stylus-Groove Relations in Phonograph Playback Process," *Acustica* **4**, 33 (1954).

THE AUTHORS



HORST REDLICH



HANS-JOACHIM KLEMP

Horst Redlich was born in 1921 and studied engineering in Berlin, Germany. He joined Telefunken in 1943 as a research engineer, and has since worked on aerials, professional tape equipment, magnetic sound recorders, and stereo records. He supervised the development of the first magnetic sound recorder in Germany for 35-mm film (in 1950, with UFA) and the first stereo records in Germany (1955-1958).

In 1951 Mr. Redlich was appointed chief engineer of the engineering branch of Telefunken-Decca (Teldec), and in 1959 he became its technical director.

Hans-Joachim Klemp was born in 1921 in Germany. From 1947 to 1951 he was a development engineer for the Tobis Company, where he worked on sound-recording equipment for motion pictures. From 1951 to 1954 he was the leader of the Defa Division's group for magnetic sound recording on film, and in 1954 he became manager of electronic engineering for Telefunken-Decca (Teldec).

Interaction of Tracing and Tracking Error*

DUANE H. COOPER[†]

University of Illinois, Urbana, Illinois

Error terms arising from the interaction of tracing and tracking errors are calculated for general stylus shapes. The terms are seen to represent an alteration in the effective shape of the stylus profile. Implications for compensation are discussed, and the possibility of the groove wall contacting the conical part of the stylus surface is analyzed in an appendix.

INTRODUCTION

IN an earlier paper,¹ a description was given of the interaction between tracing and tracking error for a playback stylus having a parabolic profile. Although it was there shown that the distortion resulting from the interaction terms would be small, some interest in the problem remains. Madsen² has proposed that the orientation of the stylus tip should be accommodated to the vertical tracking angle. Shirley³ calls attention to the argument that the symmetry of the spherically rounded tip should make it ideally indifferent to mounting orientation or to tracking angle.

Shirley is right, of course, provided the recorded waveform conforms to a rather generous peak slope limitation. Otherwise, since the spherical stylus is really only a spherical cap truncating the pointed end of a cone, the groove wall could contact the conical surface; the spherical symmetry would be spoiled; and Madsen would be right. For recordings made with responsibly monitored levels, the risk of conical contact should be very slight, and, in any case, back-facet embossing⁴ should set in before conical contact, for total included cone angles below 49° (see *Appendix*).

The question is subtler, as Shirley suspects, even when Madsen's warning is inapplicable. The indifference to orientation survives close scrutiny for the spherical tip, but not the indifference to cutting and tracking angle: there is no stylus shape which is indifferent to that. The interaction with a non-zero tracking angle causes a uniformly curved stylus tip to have the effect of one which is not uniformly curved but elliptical, as will be shown here.

This ellipticity is not to be confused with that of a stylus deliberately so shaped⁵ in an effort to minimize tracing distortion. Those tips are oriented to place the region of sharpest curvature in contact with the groove wall. The apparent elliptical shape generated by the interaction with the tracking angle presents a region of intermediate curvature to the "groove wall," the regions of sharper and blunter curvature being reached only for larger instantaneous slopes (sharper or blunter depending on the sign of the slope).

* Presented October 13, 1964, at the Sixteenth Annual Fall Convention of the Audio Engineering Society, New York.

[†] Dr. Cooper is Research Associate Professor of Physics and Electrical Engineering at the University of Illinois. His full biography appears in the January 1964 issue of the *Journal*.

In the sections that follow, a brief review of the mathematical description of tracing error is given, for which the stylus shape is assumed to have a monotonically changing slope and to make only a single-point contact with each groove wall, but be otherwise arbitrary. Without further restriction, the interaction of tracing error with tracking error is developed. So long as the tracking inclination is not zero the interaction is seen to be equivalent to an alteration in the effective shape of the stylus profile. The alteration is explicitly shown and seen to be equivalent to subjecting the shape of the profile to a skew transformation. Thus the effective shape for a circular profile becomes an ellipse. For a 15° vertical angle, the axes are tipped nearly 45°, and the curvature becomes nonuniform, departing from its nominal value by some 30 percent.

The more exact tracing compensation schemes⁶ will require the simulation of skewed stylus profiles.

GENERAL TRACING TRANSFORMATION

It has been shown^{1,7,8,9} that when the recorded waveform

$$y = y(z) \quad (1)$$

is followed by a curved stylus, a reference point on that stylus traces the waveform whose coordinates are given by the transformation equations

$$\bar{z} = z - \delta_z(s), \quad \bar{y} = y - \delta_y(s), \quad (2)$$

in which $\delta_z(s)$ and $\delta_y(s)$ are space delay and amplitude perturbations, respectively, dependent on the instantaneous slope

$$s = dy/dz \quad (3)$$

of the unperturbed waveform. Because the stylus is constrained to slide tangentially on the recorded waveform, the slope of the perturbed waveform at the reference point must match the slope of the recorded waveform at the contact point:

$$s = d\bar{y}/d\bar{z} = dy/dz. \quad (4)$$

Here, z or \bar{z} is the abscissa in the direction of the groove axis, and y or \bar{y} is an ordinate normal to that axis. These relationships are illustrated in the lower half of Fig. 1.

The functions $\delta_z(s)$, $\delta_y(s)$ also describe the shape of the

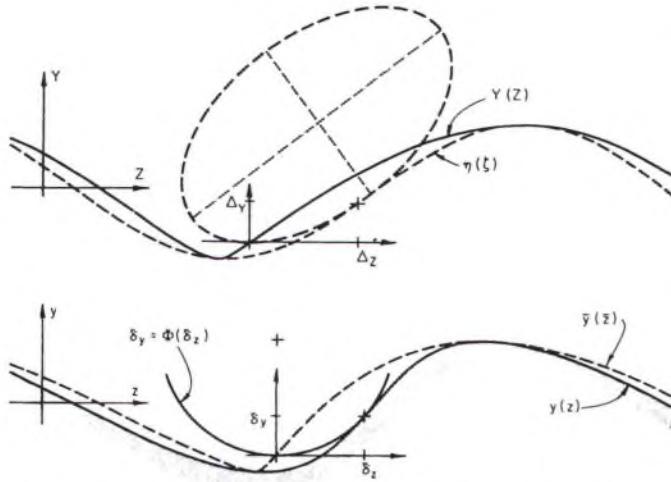


FIG. 1. Tracing and tracking transformations. In the lower half, a skewed sinusoid $y(z)$ is contacted by the circular stylus profile $\delta_y = \Phi(\delta_z)$. Successive positions of this profile cause the origin of the δ -coordinate system fixed in it to trace the distorted curve $\bar{y}(\bar{z})$, the distortion being caused by the varying delay and amplitude perturbations δ_z and δ_y . The curves plotted in the upper half are the skew images of all those in the lower. This skewing undoes that of $y(z)$ to produce the exact sinusoid $\eta(\xi)$. The tracing-distorted curve $Y(Z)$ must necessarily be interpretable as derived from $\eta(\xi)$ by way of an effective stylus shape (elliptical), the skew image of $\delta_y = \Phi(\delta_z)$.

stylus profile in a local coordinate system whose origin is the same reference point that traces the $\bar{y}(\bar{z})$ waveform. Alternatively, if the shape is explicitly given by the function

$$\delta_y = \Phi(\delta_z), \quad (5)$$

then the contact of this profile with the $y(z)$ waveform comes at the point where the slopes match:

$$s = d\Phi/d\delta_z = \Phi'(\delta_z). \quad (6)$$

At any point along the stylus profile, the profile slope is a function, as in Eq. (6), of δ_z . The inversion of Eq. (6) to express δ_z as a function of the slope and also, by way of Eq. (5), doing the same for δ_y , accomplishes the parametric description of the profile using the pair of functions, $\delta_z(s)$ and $\delta_y(s)$. The parameter is the profile slope, which, at the contact point, is also the instantaneous slope of the waveform $y(z)$ and of the waveform $\bar{y}(\bar{z})$ over at the reference point.

Equations (2) may be regarded as arising out of the integration of the pair of equations

$$z - \delta_z(s) = \bar{z}, \quad s = d\bar{y}/d\bar{z}, \quad (7)$$

because one may write

$$\bar{y}(\bar{z}) = \int d\bar{y} = \int (d\bar{y}/d\bar{z}) d\bar{z} = \int s dz - \int s d\delta_z.$$

But the two terms may be written

$$\int s dz = \int (dy/dz) dz = \int dy = y(z)$$

and

$$\int s d\delta_z = \int (d\delta_y/d\delta_z) d\delta_z = \int d\delta_y = \delta_y(s),$$

thus evaluating the second of Eqs. (2). These observations allow one to develop the point of view^{1,7} that Eqs. (7), describing the tracing transformation as one involving only a space-delay perturbation, are the more fundamental, be-

cause the appearance of the amplitude perturbation in Eq. (2) is a consequence of the delay perturbation, and does not arise independently. The pertinence of this point of view is emphasized by Eq. (7) which constitutes transformation of the slope waveform only, a transformation that makes no reference to displacement values. That transformation has a particularly simple form when the profile is parabolic, for then δ_z is proportional to s , and the transformation is the linear skew transformation.

TRACING-TRACKING INTERACTION

The purpose in this section is to follow an input waveform $\eta = \eta(\xi)$ through the transformations involved in recording and playback, schematically shown in the block diagram of Fig. 2, resulting in the output waveform $Y = Y(Z)$. The goal is to develop a basis for comparing the input with the output, so that the overall transformation may be simply represented. This is achieved by studying the transformations which the delay and displacement perturbations undergo, jointly, along with the transformations of the waveform variables. It will be seen that it is these transformed perturbations that may be used to describe the whole of the overall transformation, in the absence of net tracking error. They describe a tracing transformation for an effective stylus profile which is a skewed version of the physical profile.

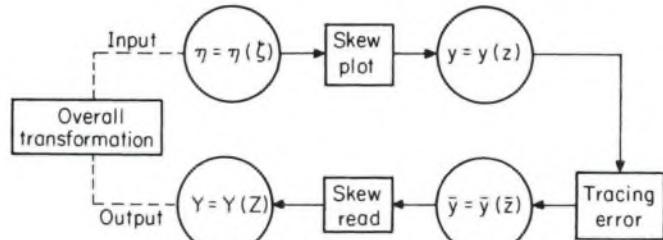


FIG. 2. Schematic representation of recording and playback transformations. The upper sequence, proceeding left to right, represents the recording transformation, with input waveform $\eta(\xi)$ and recorded waveform $y(z)$. The lower sequence, proceeding right to left, represents the playback transformation, with input $y(z)$, stylus locus $\bar{y}(\bar{z})$, and output waveform $Y(Z)$. A goal is to represent the whole by one overall transformation.

The interaction with the skew transformation of tracking¹ is more easily discussed starting from Eq. (2), because then the roles of δ_z and δ_y may be followed separately without repeating the integrations. The waveform, Eq. (1), is assumed to have been "plotted" by a cutter which would have plotted the waveform

$$\eta = \eta(\xi) \quad (8)$$

(the Greek letter ξ is the analog of z as is η the analog of y), except that it plots in skew coordinates of inclination K (the tangent of the skew angle), i.e., it plots

$$z = \xi - K\eta, \quad y = \eta, \quad (9)$$

as representing the skew transformation T_K . Equations (9) are to be substituted into Eq. (2), using the slope equations

$$s = dy/dz = d\bar{y}/d\bar{z} = (d\eta/d\xi)/(1 - Kd\eta/d\xi),$$

or

$$s = \sigma/(1 - K\sigma), \quad (10)$$

for

$$\sigma = d\eta/d\zeta.$$

This being understood, the mention of σ may be suppressed for the time being, to write the overall transformation up to this point compactly as

$$\bar{z} = \zeta - K\eta - \delta_z, \quad \bar{y} = \eta - \delta_y. \quad (11)$$

Upon playback with a pickup of inclination κ the skew transformation $T_{-\kappa}$ is imposed. It is

$$Z = \bar{z} + \kappa \bar{y}, \quad Y = \bar{y} \quad (12)$$

so that one has for the overall transformation

$$Z = \zeta - K\eta - \delta_z + \kappa\eta - \kappa\delta_y, \quad Y = \eta - \delta_y. \quad (13)$$

To introduce a more compact notation for showing the explicit role of slopes, substitution in Eq. (10) is to be represented by

$$\begin{aligned} \delta_z(s) &= \delta_z(\sigma), & \delta_y(s) &= \delta_y(\sigma), \\ s &= \sigma/(1-K\sigma). \end{aligned} \quad (14)$$

Then Eq. (13) becomes

$$\begin{aligned} Z &= \zeta - (K - \kappa)\eta - [\delta_z(\sigma) + \kappa\delta_y(\sigma)], \\ Y &= \eta - \delta_y(\sigma). \end{aligned} \quad (15)$$

A further reduction in form may be obtained. The overall slope transformation is

$$S = dY/dZ = (d\bar{y}/d\bar{z})/(1 + \kappa d\bar{y}/d\bar{z}),$$

which is seen, with the help of Eq. (10) and a little manipulation, to be

$$S = \sigma/[1 - (K - \kappa)\sigma]. \quad (16)$$

Thus, it is seen that if the tracking error be made to vanish, i.e., if $K = \kappa$, then one may substitute S for σ everywhere, and the overall transformation becomes the tracing transformation

$$Z = \zeta - \Delta_Z(S), \quad Y = \eta - \Delta_Y(S), \quad (17)$$

in which the tracing error itself obeys the transformation

$$\begin{aligned} \Delta_Z(S) &= \delta_z(S) + \kappa\delta_y(S), \\ \Delta_Y(S) &= \delta_y(S). \end{aligned} \quad (18)$$

The interpretation of Eq. (18) may be made quite explicit. Equations (14) show that the functions δ_z, δ_y are numerically equal to the functions δ_s, δ_y , but differ only in the form of the parameter upon which they depend. Upon carrying through the elimination of the parameter, therefore, one must have δ_y being the same function of δ_z

$$\delta_y = \Phi(\delta_z),$$

as described the dependence of δ_y upon δ_z in Eq. (5). The same is not true, however, of Δ_Z and Δ_Y . Upon carrying through the elimination one sees that the coordinates in which the stylus profile is to be described have suffered the skew transformation

$$\Delta_Z = \delta_z + \kappa\delta_y, \quad \Delta_Y = \delta_y, \quad (19)$$

and that from Eq. (17) the effective shape of the stylus profile has been altered by the playback tracking transformation $T_{-\kappa}$.

EFFECTIVE STYLUS SHAPE

The skew transformation is a subset of the projective transformations.¹⁰ As such, it transforms conic sections into conic sections. Among its other limitations, it preserves connectivity in the finite part of the plane. In particular, it transforms straight lines into straight lines of altered slope; it transforms parabolas into parabolas, circles into ellipses, etc. A circular stylus profile has the effective shape, then, of an elliptical profile because of Eq. (19).

To investigate the details it is convenient to abandon the cumbersome notation used above, since, in this section, there will be no need to keep track of alterations in waveforms in parallel with those of stylus shape. Let the circle be the prototype stylus profile. When the origin of the coordinates ζ, η is placed at the center, a circle of radius c is described by the equation

$$\zeta^2 + \eta^2 = c^2. \quad (20)$$

The transformation (substitution)

$$\zeta = \bar{z} - \kappa \bar{y}, \quad \eta = \bar{y} \quad (21)$$

results in the equation

$$\bar{z}^2 + (1 + \kappa^2)\bar{y}^2 - 2\kappa\bar{z}\bar{y} = c^2, \quad (22)$$

which describes a conic section with tilted axes, as betrayed by the non-vanishing coefficient of the cross-term, $\bar{z}\bar{y}$.

One may determine the angle of tilt by substituting the rotation equations

$$\begin{aligned} \bar{z} &= z \cos \theta - y \sin \theta, \\ \bar{y} &= z \sin \theta + y \cos \theta \end{aligned} \quad (23)$$

into Eq. (22) and selecting the value of θ to force the new coefficient of the cross-term zy to vanish. When this is done the equation in z and y reduces to

$$(z/a)^2 + (y/b)^2 = 1. \quad (24)$$

This is achieved by setting

$$\theta = \frac{1}{2} \tan^{-1}(1/\kappa), \quad (25)$$

which, for small κ , places θ very near to 45° , differing from that value by approximately $\kappa/2$ radians. Equation (24) is the standard one for an ellipse of major and minor axes $2a$ and $2b$. The values a, b are functions, via Eq. (25), of κ and c , and they also give the maximum and minimum radii of curvature of the ellipse. These radii are a^2/b and b^2/a .

For a vertical angle of 15° the vertical component of κ is 0.268, but the component for each of the groove walls is $0.268/\sqrt{2}$. For this reduced value of κ , the extreme values of curvature for the effective stylus profile differ from the c -value by some 30%, being at one part blunter and at another part sharper.

DISCUSSION

In the upper part of Fig. 1 the overall transformation to $Y(Z)$ from $\eta(\zeta)$ is depicted. It recapitulates the analysis in the foregoing sections, in showing that when the skewing of $y(z)$ is undone in playback, the image of the distorted curve $\bar{y}(\bar{z})$ becomes $Y(Z)$ which differs from the image $\eta(\zeta)$ of $y(z)$ by the tracing errors, Δ_Y and Δ_Z . Because of the preservation of connectivity, those errors must be gen-

erated by a profile which is the image of $\delta_y = \Phi(\delta_x)$. That image is the elliptical form shown. For the purposes of clarity, the skew angle was exaggerated to 30° in the constructions of the figure.

As calculated above, the alterations in curvature are not nearly so great as in the figure, for practical cases. Also, the regions of extreme departures in curvature from the nominal c -value are attained only for instantaneous waveform slopes approaching unity. Thus, any extra distortion arising from the curvature alterations should be undetectable in the presence of normal amounts of tracing distortion. Again it should be pointed out that the well known observation¹¹ that tracing distortion arises independently in the two signals at each of the two groove walls, is not prejudiced by the properties of the tracing-tracking interaction, since only an alteration in effective stylus shape is involved.

In seeking to remove tracing distortion by a compensatory pre-shaping of the groove wall modulation,¹² an observable part of the remaining distortion may be due to this interaction. The fault is that the simulated stylus used in the compensation is used to generate a waveform that is skew-plotted by the cutter. Upon unskewing in playback, the effective stylus shape does not match the simulated stylus shape. To put it another way, the simulated stylus shape should be skewed, so that when the waveform is skew-plotted by the cutter the image of the simulated profile will agree with the shape of the physical stylus profile actually used in playback. Thus, some of the insensitivity in the residual distortion to variations in stylus curvature, as observed,¹³ may be removed and the residual distortion reduced.

APPENDIX

The condition for conical contact mentioned in the *Introduction* may be deduced from a consideration of the pinch effect in playback of a lateral (mono) recording, and a generalization to the stereo case. The pinch effect narrows the quiescent angle θ_0 between groove walls to the angle θ , measured in a plane normal to the instantaneous groove axis. The relation is

$$\tan(\theta/2) \cos a_m = \tan(\theta_0/2), \quad (26)$$

in which $\tan a_m$ would be the instantaneous slope of the mono modulation. The interpretation is that, for $\tan a_m$ representing a peak value, conical contact would just be avoided for a total included cone angle equal to θ .

This interpretation is valid for each groove wall separately, but the slope of the wall modulation is

$$\tan a = \sqrt{2} \tan a_m$$

for $\theta_0 = 90^\circ$. With these substitutions, one may reduce Eq. (26) with the help of half-angle formulae, concluding that the peak wall-modulation slope would be

$$s_{max} = \tan a = \sqrt{[\cos \theta / (1 - \cos \theta)]}, \quad (27)$$

for the avoidance of conical contact on the part of a single wall of a stereo groove. It is seen, for example, that a total included cone angle of 60° would provide clearance for peak wall slopes up to unity.

It is true that greater slopes may be cut. Bauer points out¹⁴ that back-facet embossing limits only the smaller of

the two slopes of opposing signs that result if skew cutting is done. If the dihedral angle between the cutting facet and the back facet is 45° , then the back slope may be, from Eq. (10),

$$1 = \delta_{max} / (1 + K\sigma_{max}),$$

at most. The steeper front slope is

$$s_{max} = \sigma_{max} / (1 - K\sigma_{max}),$$

where σ_{max} is the peak slope of the modulation that would have obtained if the cutting inclination, K , were zero. Determining the value of σ_{max} from the back-slope limitation one has, after a little reduction,

$$s_{max} = 1 / (1 - 2K). \quad (28)$$

For a vertical cutting angle of 15° , one has $K = 0.268/\sqrt{2}$. From this one finds $s_{max} = 1.39$. Solving Eq. (27) for θ , one has

$$\cos \theta_{max} = s_{max}^2 / (1 + s_{max}^2), \quad (29)$$

which yields, for the above value of s_{max} , the value $\theta_{max} = 49^\circ$. This may be compared with the NAB standard¹⁵, requiring θ to fall in the range 40° to 55° .

The 55° figure would allow a 5° margin (to allow for an orientation error, presumably) only for a modulation level such that the peak slope (taken with respect to a coordinate system in which the quiescent walls were normal to the displacement axis) would never exceed unity. Ordinarily, this would be thought to be a generous limitation on slope. Providing a similar margin under Bauer's even more generous limitation would require a total included cone angle of below 44° , or else require inclining the cone axis as Madsen actually proposed.¹⁶

REFERENCES

- ¹ D. H. Cooper, "Integrated Treatment of Tracing and Tracking Error," *J. Audio Eng. Soc.* 12, 2 (1964).
- ² E. R. Madsen, "Vertical Tracking Angle—a Source of IM Distortion," *Audio* 46, 21 (November, 1962).
- ³ G. Shirley, "Cutting Remarks from a Stylus Maker," *Audio* 48, 10(L) (June, 1964).
- ⁴ C. R. Bastiaans, "Further Thoughts on Geometric Conditions in the Cutting and Playing of Stereo Disks," *J. Audio Eng. Soc.* 11, 6 (1963).
- ⁵ J. Kogen and R. Samson, "The Elliptical Stylus," *Audio* 48, 33 (May, 1964).
- ⁶ E. C. Fox and J. G. Woodward, "Tracing Distortion—Its Cause and Correction in Stereodisk Recording Systems," *J. Audio Eng. Soc.* 12, 2 (1964).
- ⁷ D. H. Cooper, "Unified Analysis of Tracing and Tracking Error," *IEEE Trans. on Audio AU-11*, 141 (1963).
- ⁸ M. S. Corrington, "Tracing Distortion in Phonograph Records," *RCA Rev.* 10, 241 (1949).
- ⁹ W. D. Lewis and F. V. Hunt, "A Theory of Tracing Distortion in Sound Reproduction from Phonograph Records," *J. Acoust. Soc. Am.* 12, 348 (1941).
- ¹⁰ See, for example, W. C. Graustein, *Introduction to Higher Geometry* (The Macmillan Company, New York, 1952).
- ¹¹ M. S. Corrington and T. Murakami, "Tracing Distortion in Stereophonic Disc Recording," *RCA Rev.* 19, 216 (1958).
- ¹² E. C. Fox and J. G. Woodward, *op. cit.*
- ¹³ J. G. Woodward, "Reducing Distortion in Stereo Phonograph Systems," *Audio* 48, 23 (February, 1964).
- ¹⁴ B. B. Bauer, "15-Degree Vertical Angle—A Key to Better Stereo Sound," *Audio* 47, 21 (November, 1963).
- ¹⁵ J. J. Bubbers, "A Report on the Proposed NAB Disc Playback Standard," *J. Audio Eng. Soc.* 12, 51 (1964).
- ¹⁶ E. R. Madsen, *op. cit.*

Construction of Tracing Correlator Waveforms

DUANE H. COOPER

University of Illinois, Urbana, Illinois

An early conception of a tracing correction scheme, now embodied in the RCA Dynamic Recording Correlator, is discussed to show how it leads to methods for giving a graphical analysis of correlator error. An example is worked through and tends to show the error to be remarkably small.

INTRODUCTION

ON November 8, 1946, B. M. Oliver prepared a Bell Telephone Laboratories Technical Memorandum describing a corrector circuit for eliminating tracing distortion in vertical-cut disc recordings.¹ The scheme was virtually identical with that brought to a practical level of development at RCA as the Dynamic Recording Correlator,² some fifteen years later. There appears to have been no publication of Oliver's scheme in the open literature, however, and no evidence that it had any influence on subsequent developments.

Apart from noting an interesting historical curiosity, there would be little point in calling attention to Oliver's scheme, except that his description of the operating principles of the correlator is somewhat novel. Indeed, his description is remarkably lucid and appealing, and it suggests a simple way to make a graphical construction of the waveforms produced by the correlator. In making these constructions, it is possible to obtain an intuitive appreciation for the style of approximation being employed.

WAVEFORM CONSTRUCTION

Oliver observed that in playback the stylus moves as a rigid body, so that every point in the stylus traces the same curve as every other point, except the curves differ from one another by being translated by the amount of the separation of one stylus point from another. In particular, if one thinks of the record groove walls as stationary and regards the stylus as sliding in contact with these walls, then the curve traced by any point fixed in the stylus is the output waveform in space coordinates, as affected by tracing error. Further, it is sufficient to consider only points

along the stylus profile in following the stylus motion. (The stylus profile is the locus of all points in the stylus tip that make contact with a given groove wall.) Figure 1

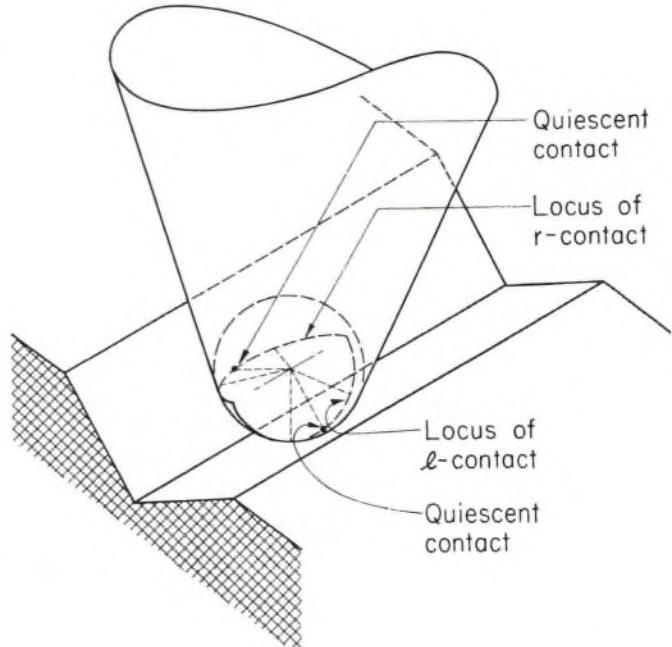


FIG. 1. Stylus-groove geometry. The conical stylus shank is shown with its tip rounded to a spherical shape. The points of contact with two groove walls are shown for an unmodulated (quiescent) groove. When modulated, the groove walls will be sloped so that the contacts will move to new positions along the loci of contact. These loci are fixed in the tip, lie in planes normal to the quiescent wall but parallel to the quiescent groove axis, and serve as the stylus profile for each wall, in discussing tracing error.

shows the stylus-groove-wall relationship and an indication of this profile.

To see the effect of tracing error, Oliver appeals to the idea of the envelope of a family of curves. This idea is seen with the help of Fig. 2. The five points marked lie on the

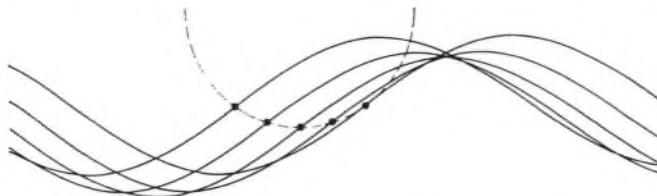


FIG. 2. Stylus trajectory family. The five curves show the trajectories of five points fixed in the playback stylus profile, for sinusoidal motion of the stylus. The lower envelope of the complete family would represent the wall modulation needed to produce such motion.

circumference of a circle, shown dashed, which is intended to represent the stylus profile. The stylus is assumed to follow a sinusoidal trajectory, so that each of the points on the profile traces a sinusoidal trajectory. These are the sinusoids shown. They are identical to one another except for being translated along the direction of the groove axis by the amount of the point separations, and displaced vertically (i.e., in the direction of the wall normal) by the amount of vertical separation between points. The lower envelope of this family of trajectories is the curve which follows whichever trajectory lies lowest at any point. Since only five points on the profile have been used, a complete family has not been shown, and the lower envelope evident here is only an approximation to the envelope of the complete family. Obviously, however, the inclusion of curves traced by points within the profile would not contribute to the determination of the envelope.

The lower envelope of the complete family of stylus trajectories must then be the full space locus traced by the contact between the groove wall and the stylus tip. It is the tracing of this curve, the lower envelope, which governs the shape of the trajectory curves, any one of which, as stated, represents the output waveform. Thus, if one assumes that stylus-groove-wall contact is unfailingly maintained and that the groove wall is sufficiently stiff (standard assumptions for treating the purely geometric aspects), then one has in this construction a prescription for determining the groove-wall waveform so that the desired output waveform shall be faithfully produced.

Though the construction has been illustrated only for the determination of the wall waveform needed to produce a sinusoidal output, assuming a circular profile, the method is perfectly general. Any output waveform whose curvature at any point is not sharper than that of the stylus profile could have been assumed. Also, any profile shape, always curving in the same direction, could have been assumed. The beauty of the construction is that it makes opportunities for mechanizing the formation of the wall waveform immediately evident.

Following Oliver, mechanization requires that distances along the groove axis be simulated by time intervals, so that the translations in that direction among the members

of the trajectory family may be simulated by fixed time delays with the help of a tapped delay line. Further, the displacements normal to the groove wall among the members of the trajectory family may be simulated by a schedule of fixed biases for the taps of the delay line. When these two conditions are met, a time simulation of the trajectory family is generated, with a different member of the family being presented at each tap. The lower envelope is then constructed by the operation of a switching network to transmit the waveform which, with bias, lies lowest at any instant, until some other member of the waveform family seizes control by presenting lower lying values. The switching network then selects that tap and transmits its waveform values, with bias, etc.

It is seen that the block diagram of Fig. 3 is appropriate

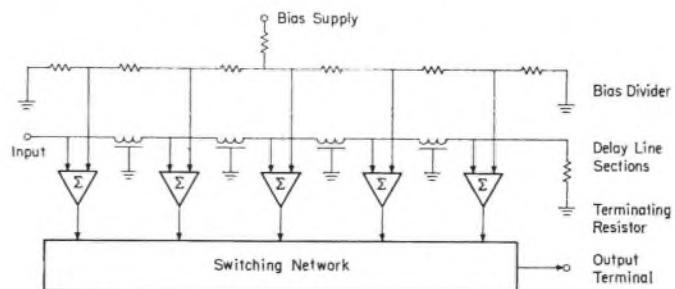


FIG. 3. Recording correlator. The block diagram shows the operations needed to construct a time simulation of the lower envelope of the finite trajectory family of Fig. 2, except that the details of the switching network are not shown. Also not shown are the details of accommodating the simulation to the groove speed.

to the performance of these functions. It is not the purpose here to describe the realization of the switching network, so that those details have been omitted from the figure. A variety of arrangements may be proposed, however; Oliver suggested a simple diode-resistor network which also included the biasing arrangement. Also, only five taps are shown, so that an exact construction of the lower envelope is not provided. A better approximation would be available if more taps were used; Oliver supposed that some 11 to 15 taps would be appropriate. An odd number should be used to minimize the need for switching, in low-slope waveforms.

Finally, the interpretation of translations along the groove axis as time delays involves the groove speed, which varies with cutting radius. To accommodate this variation, one may divide the radius into zones within which the groove speed is nearly constant, and arrange to switch to a new set of delay taps when the radius enters a new zone. The biases would remain unaltered. Alternatively, if a sufficient number of taps were available, it would suffice to adjust the bias supply to accommodate the variation in radius, and use a fixed set of taps. This would result in a coarser approximation of the stylus profile at the larger cutting radii, but this seems reasonable in any case, since only a slighter correction need be invoked there. Oliver proposed the latter accommodation to the groove speed.

The diagram of Fig. 3 is close enough to those published for the Dynamic Recording Correlator² to show that Oliver

succeeded in describing the essential features of the correlator in 1946. The RCA Correlator uses a dozen taps at each cutting-radius zone and switches to another set upon entering a new zone. It is necessary to invoke the correlator action for the correction of each wall waveform separately, so that a practical device will use two identical but independent channels, even in making mono (i.e., purely lateral-cut) recordings. Only for Oliver's original application, vertical-cut recording, would a single channel suffice.

CORRELATOR APPROXIMATIONS

The correlator uses only a finite number of points to represent the stylus profile, and the envelope is constructed from only a finite family of stylus trajectories, as in Fig. 2, where five are shown. Close inspection of Fig. 2 will show that the envelope of the finite family is a continuous curve, defined in sections, but that there is a jump in slope in going from one section to the next. This fact is more evident in Fig. 4 showing a plot of envelope slope values. (The upper part of this figure is a reproduction of Fig. 2.)

These discontinuities in slope will always be less evident in the finite envelope than in the corresponding slope waveform itself. However, for reproducing sinusoidal waveforms that are less sharply curved compared to the stylus profile, the slope discontinuities will appear as smaller jumps than

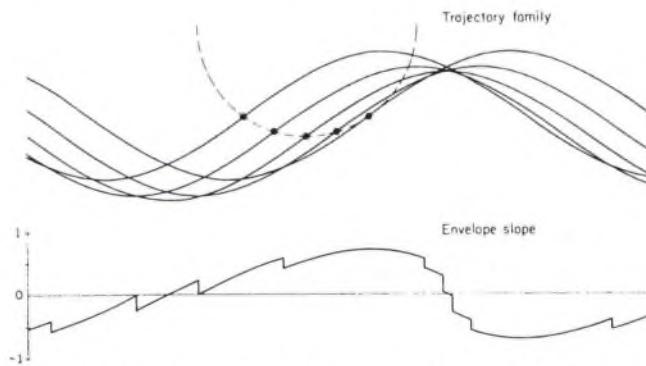


FIG. 4. Comparison of a trajectory envelope with the envelope slope. The gently scalloped structure of the finite envelope makes its appearance as discontinuities in the slope waveform, marking the unequally spaced switching points. The approximation used by the correlator does not suggest any simple mathematical characterization.

shown here. Further, the total number of switching points per cycle will depend on the total range of slope values, so that signals of sufficiently low level, on a peak slope (or peak velocity) scale, may invoke no switching whatever.

One naturally wonders if some kind of linear filtering can smooth out these jumps to lead to a better approximation. The answer is "yes," of course, but it is an affirmative answer that must be carefully qualified. The slope waveform of Fig. 4, for example, would certainly be improved by transmission through a low-pass filter whose impulse response were comparable in width to the spacing between steps. The trouble is that the spacing between steps is not uniform. Further, the number of steps per cycle is the same for waveforms of the same peak slope but varying frequency, so that the average step spacing

depends on the frequency. Thus, while a filter with a fixed impulse response, implying a fixed-pass band, would offer some improvement in some instances, the improvement would be negligible in others. This is, of course, the limited improvement that would be offered by the finite response speed of the cutting mechanism. On the other hand, the development of a filter that could adapt itself to the varying step spacing would be an achievement surpassing that of the correlator itself.

It should also be clear that no kind of linear interpolation among the family of curves in the upper part of Fig. 4 can lead to a closer approximation to the complete envelope. The reason is that the complete envelope always lies as low as, or lower than, the finite envelope. Indeed, the finite envelope may be described as a "scalloped" version of the complete envelope such that at the "points" of the scallops it lies above the complete envelope, touching the latter at only one place between points. Some kind of extrapolation would be needed to reach nearer the complete envelope.

Such extrapolation often occurs if the finite knee curvature of the diode characteristic curve is properly exploited in the switching. The effect is to round down the points of the scallops, so that in the slope curve the discontinuities in slope are replaced by gentler ramp-like sections. This is a refinement one would certainly want to include in a more ambitious treatment. Leaving it out here places the present treatment in the "worst-case" category, as does using an unrealistically small number of taps. Adherence to the present simple-minded analysis has the merit of clarity, as well as the virtues of a worst-case treatment. The worst-case distortion, shortly to be estimated, will, in fact, be seen to be surprisingly small.

While the constructions of Fig. 4 do lead to intuitive insights into the nature of the approximation, they do not seem to lead to any simple mathematical description. The idea of switching, for example, seems to suggest a treatment by means of the sampling theorem. However, the fact that the switching points are unequally spaced, together with the fact that variously delayed sections of the waveform are transmitted between switching points, seems to remove the analysis from the domain of the classical theory.³ This should not be too surprising, since that theory is essentially one describing linear processes.

APPROXIMATIONS SEEN IN PLAYBACK

Of more pressing interest, however, than the characterization of the style of approximation inherent in the correlator waveforms, is the nature of the approximation resulting for the playback waveforms. The playback waveform representing stylus displacement can also be constructed using envelope ideas, following Oliver. Since we want an excellent approximation so that the correlator style of approximation will not be obscured, the family should have as many members as convenience will allow. Equivalent to Oliver's construction is the construction of a family of circles (for a circular playback-stylus profile) with centers spaced along the curve to be traced; then the upper envelope of this family is the result of the tracing. This construction was

done upon the correlator displacement waveform of Fig. 4, using a family of 200 circles, with the top curve of Fig. 5 showing the result.

The disagreement with the sinusoidal shape is almost indiscernible, with only a few small "wrinkles" evident near the minimum. This is impressive evidence indeed that the approximation, whatever its characterization, is excellent.

The appraisal of the approximation upon playback is more radically altered upon constructing the slope wave-

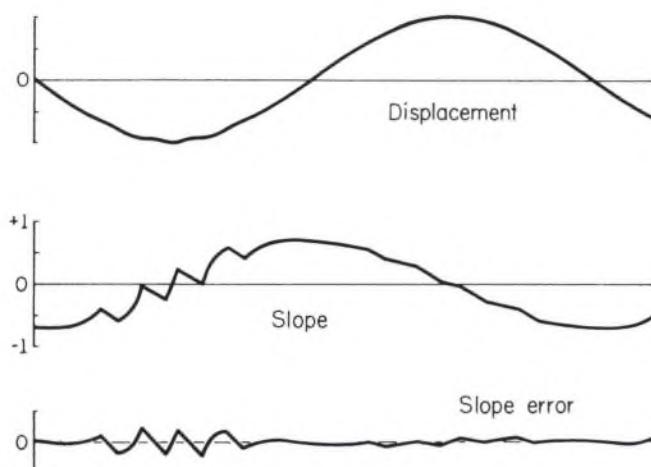


FIG. 5. Tracing of transformed correlator waveforms. The top waveform is the result of applying the tracing transformation to the envelope of the finite trajectory family of Fig. 4. The middle waveform results from applying a skew transformation to the slope waveform, as the correct representation of the tracing transformation on a slope basis. The bottom waveform is the error in the slope waveform. Since a stylus profile matching that in Fig. 4 was used, only small departures from sinusoidal shape appear.

form. This construction is achieved by applying a suitable skew transformation^{4,5} to the lower waveform of Fig. 4. The discontinuities in slope have been transformed to gentler "ramp" sections in the curve, the resulting curve showing a much smoother appearance. Agreement with a sinusoidal shape is quite close, as shown in the bottom curve of Fig. 5, the residual distortion remaining after subtracting the undistorted sinusoid.

The principal effect seems to be a short burst of oscillation at about the 13th harmonic of the wave shown, with a peak strength of about 20% on an equivalent sinewave basis. The near imperceptibility of its appearance in the displacement waveform is readily understood, for upon integration its relative intensity should be diminished by the factor $\frac{1}{13}$ to be a little less than 2%. These "squeaks" in the velocity waveform last for about a fifth of a period, so that on a wave analyzer the spectral components in the neighborhood of the 13th harmonic would have a strength of about 4% altogether. On a displacement basis, some 0.3% would be revealed.

DISCUSSION

This graphical method of analysis, while simple and intuitively accessible, shares a weakness in common with numerical methods in failing to reveal the comprehensive

logical structure one expects to see in an analytic formulation. Actually, the method does seem to have a slight edge over blind numerical analysis. One can see how "squeaks" arise, and how their amplitude is related to peak curvature and how the harmonic number is related to the number of switching actions per cycle, and thus to peak slope. Though these ideas are not precise, they do have quantitative substance, allowing even an estimate of the way the distortion would be reduced if more taps were used which conforms to actual practice.

The failing in comprehensiveness is more telling in another way, however. The limited success in predicting harmonic content does not carry over to the prediction of intermodulation distortion. In dealing with a memory-coupled nonlinearity, as experience with undiluted tracing and tracking distortion shows, intermodulation can produce spectral components of a much greater intensity relative to the harmonic content than would have been expected if a memoryless nonlinearity had been the governing one.⁶ It would be reckless, then, to try to predict the intermodulation components that the present residual nonlinearity might generate.

The availability of high-speed computing equipment does offer a partial solution to the problem of comprehensiveness, by virtue of the ability it offers to simulate a large variety of particular cases, providing at the same time a measure of the distortion each presents. In this way, trends can be seen that would be too laborious to explore by hand methods. Intermodulation analysis is a good case in point as one involving excessive labor by hand. If a little recklessness may be allowed, however, one might venture to predict that machine analysis would not show nearly as much residual intermodulation as has been measured,² and that the experimental data will be seen as evidence, largely, of other sources of distortion.⁶

One of the more familiar of these "other sources" is groove-wall deformation. This has, of course, a large linear component which only serves to alter the memory coupling of the tracing nonlinearity. The nonlinear component generates distortion on its own account. These effects have been systematically left out of the present discussion, particularly in the constructions of Fig. 5. One may guess, again with some degree of recklessness, that the linear component will tend to alter the strength of the harmonic components already observed, enhancing them if they fall at the wall-stylus resonance or diminishing them if they fall above, but otherwise having little effect. Further, one would suppose that the nonlinear component would generate additional distortion components but not materially alter those components already found. Though the theory is still not one which allows a clear view of its consequences, requiring power series expansions for its manipulation, it also does not seem to offer a definite explanation of the experimental observations of residual distortion.

Whatever the source of the residual distortion, its exact identification is of little importance if it is not to be unmasked by the removal of tracing and tracking distortion. Means for doing this are now available to the industry. A first step in making the identification is establishing whether

the residual arises largely from imperfections in the means for removing these known distortions. The present graphical analysis tends to show that the action of the correlator cannot be blamed for the observed residual distortion. More comprehensive numerical analyses are needed, however.

REFERENCES

1. B. M. Oliver, Hewlett-Packard Company, private communication; March 17, 1964.
2. E. C. Fox and J. G. Woodward, "Tracing Distortion—Its Cause and Correction in Stereodisk Recording Systems," *J. Audio Eng. Soc.* 11, 294 (1963).
3. B. M. Oliver, J. R. Pierce, and C. E. Shannon, "The Philosophy of PCM," *Proc. IRE* 36, 1324 (1948).
4. D. H. Cooper, "Unified Analysis of Tracing and Tracking Error," *IEEE Trans. on Audio AU-11*, 141 (July-Aug. 1963).
5. D. H. Cooper, "Integrated Treatment of Tracing and Tracking Error," *J. Audio Eng. Soc.* 12, 2 (1964).
6. D. H. Cooper, "On Tracking and Tracing Error Measurements," *J. Audio Eng. Soc.* 12, 312 (1964).

THE AUTHOR



Duane H. Cooper was born in Gibson City, Illinois, on August 21, 1923. He received the B.S. degree in 1950 and the Ph.D. degree in 1955, both from the California Institute of Technology, Pasadena.

Since 1954 he has been with the Coordinated Science Laboratory (then the Control Systems Laboratory) of the University of Illinois, Urbana, where he is Research Associate Professor of Physics and Electrical Engineering. In recent years his principal interests have been statistical detection, information and decision theory, with applications to pattern recognition, and the applications of digital computing systems in instrumentation and simulation of physical systems by Monte Carlo methods.

Dr. Cooper is a member of the American Physical Society and of the Audio Engineering Society.

LETTERS TO THE EDITOR

Note to Members: This is your column. It is designed for the discussion of papers published in the Journal and other pertinent topics about which you feel strongly.

MISINTERPRETATION OF VERTICAL TRACING ERROR

DUANE H. COOPER

University of Illinois, Urbana, Illinois

Vertical tracing error was given its first analytical description in the 1941 paper of Lewis and Hunt,¹ but for the hill-and-dale recording technique in use at that time. There is, however, a significant difference between that modulation and the vertical modulation component of present-day stereo practice. The difference is easily overlooked, and has led Redlich and Klemp into unfortunate numerical errors in a recent paper.² The purpose of this letter is to correct those errors, and to set the record straight for the benefit of other investigators.

Lewis and Hunt use the following sentence in their *Introduction*: "As in the previous references, it is assumed throughout that . . . the stylus is always in . . . contact with . . . the bottom . . . of the record groove." Thus, in analyzing tracing distortion for the vertical cut, they supposed that a single wall—the bottom wall—would be modulated. Only in considering lateral cut did they analyze the effect of two-wall modulation. In this way, they found the cubic motion in the lateral to be half that which would obtain for the vertical. Present-day stereo practice, however, modulates a V-shaped groove (both walls) for the vertical component. For this practice, using the 45°-45° orientation of walls, the cubic motion for vertical modulation alone is equal to that which obtains for lateral alone, considering only tracing error.

Corrington³ extended the analysis of Lewis and Hunt, but followed the same assumptions, so that his vertical-cut case was also for the old-fashioned bottom-wall hill-and-dale modulation. Since it is just the case for single-wall modulation, however, these early results do apply to today's practice in the instance in which only one of the two walls

would be modulated. It follows from the fact that tracing error is generated independently at each wall,⁴ that the application of these early results to modern practice is a simple matter of interpretation.

If, for example, Eq. (8) of Lewis and Hunt be recast in terms of velocity values alone, for a parabolic stylus profile, then it may be interpreted as pertaining to the right stereo wall. Then it is written as

$$V_R(t) = \frac{c}{2v_y^2} \frac{d}{dt} v_R^2(t) + \frac{c^2}{3!v_y^4} \frac{d^2}{dt^2} v_R^3(t) + \dots \quad (1)$$

The same form connects $V_L(t)$ with $v_L(t)$ for the left wall, except for a reversal of the sign of c , the stylus radius of curvature. Here V refers to the output stylus motion and v refers to the wall motion; v_y is the groove speed. Writing

$$v_R = \frac{1}{2}(v_x + v_y)\sqrt{2}, \quad v_L = \frac{1}{2}(v_x - v_y)\sqrt{2}, \quad (2)$$

and

$$V_x = \frac{1}{2}(V_R + V_L)\sqrt{2}, \quad V_y = \frac{1}{2}(V_R - V_L)\sqrt{2}, \quad (3)$$

one can use Eq. (1) and its mate to write the expansion for the vertical component of stylus motion,

$$V_y = v_y + \frac{\bar{c}}{2v_y^2} \frac{d}{dt} (v_y^2 + v_x^2) + \frac{\bar{c}^2}{3!v_y^4} \frac{d^2}{dt^2} (v_y^3 + 3v_y v_x^2) + \dots \quad (4)$$

and the expansion for the lateral component,

$$V_x = v_x + \frac{\bar{c}}{2v_y^2} \frac{d}{dt} (2v_x v_y) + \frac{\bar{c}^2}{3!v_y^4} \frac{d^2}{dt^2} (v_x^3 + 3v_x v_y^2) + \dots \quad (5)$$

In Eqs. (4) and (5) there is written

$$\bar{c} = \frac{1}{2}c\sqrt{2}, \quad (6)$$

¹ W. D. Lewis and F. V. Hunt, "A Theory of Tracing Distortion in Sound Reproduction from Phonograph Records," *J. Acoust. Soc. Am.* 12, 348 (1941).

² H. Redlich and H. J. Klemp, "A New Method of Disc Recording for Reproduction with Reduced Distortion: the Tracing Simulator," *J. Audio Eng. Soc.* 13, 111 (April, 1965).

³ M. S. Corrington, "Tracing Distortion in Phonograph Records," *RCA Rev.* 10, 241 (1949).

⁴ M. S. Corrington and T. Murakami, "Tracing Distortion in Stereophonic Disc Recording," *RCA Rev.* 19, 216 (1958).

as representing a reduced stylus radius; this follows naturally from Eqs. (2) and (3) and does not have to be separately introduced. The factor $\frac{1}{2}\sqrt{2}$ corresponds to the $\cos\gamma$ factor of Lewis and Hunt. It happens also to agree with the $\sin\gamma$ factor which they would have found if they had treated vertical modulation for the two-wall case.

The term $3v_y v_x^2$ in Eq. (4) vanishes for pure vertical modulation. Lewis and Hunt do not obtain it, of course. Also, the terms $2v_x v_y$ and $3v_x v_y^2$ in Eq. (5) vanish for pure lateral modulation and, again of course, are not obtained by Lewis and Hunt. Unfortunately, Redlich and Klemp do not appear to be aware of these terms, though they are needed for general directions of modulation. Thus, the plots in the right-hand column of their Fig. 4 are incorrect. Correct plots are given in this letter (see Fig. 1), although it is not sufficient to speak of single directions of modulation. Some frequencies in the program material may fall in one direction while other frequencies may simultaneously fall in other directions, producing crossmodulation tones⁵ falling in yet other directions.

For a single direction of modulation, making an angle θ with respect to the lateral, however, one sets

$$v_y = v \sin \theta, \quad v_x = v \cos \theta,$$

to make the calculations leading to Fig. 1. The quadratic term in V_y is independent of θ , but the corresponding one in V_x is proportional to $\sin 2\theta$. Also, the cubic term in V_y is proportional to $\frac{1}{2}(3 \sin \theta + \sin 3\theta)$ and that in V_x is proportional to $\frac{1}{2}(3 \cos \theta - \cos 3\theta)$, as it turns out after some reduction.

These y -coefficients are plotted vs the x -coefficients for various values of θ to make the vector plots of Fig. 1. For the linear term, the four quadrants are numbered so that the four-fold occupation of the single quadrant for the quadratic term may be indicated. For the cubic term, there is a one-for-one correspondence of quadrants, with the linear, so that the quadrant marking is unnecessary. The cubic 30° vector lies so close to the one at 45° that it was omitted for clarity, though the point on the curve where it would fall is indicated. An extra 5° vector is shown in the cubic plot. The cubic curve is a hypocycloid, the locus traced by a point on the circumference of a circle of unit diameter as it rolls on the inside of another circle whose diameter is four units; it also obeys

$$(x+y)^{2/3} + (x-y)^{2/3} = 2,$$

which is the equation of a curve called the *astroid*.

It is to be hoped that Redlich and Klemp will be able

⁵ D. H. Cooper, "On Tracking and Tracing Error Measurements," *J. Audio Eng. Soc.* 12, 312 (1964).

Linear term
 $x = \cos \theta$
 $y = \sin \theta$

Quadratic term
 $x = \sin 2\theta$
 $y = 1$

Cubic term
 $x = \frac{1}{2}(3 \cos \theta - \cos 3\theta)$
 $y = \frac{1}{2}(3 \sin \theta + \sin 3\theta)$

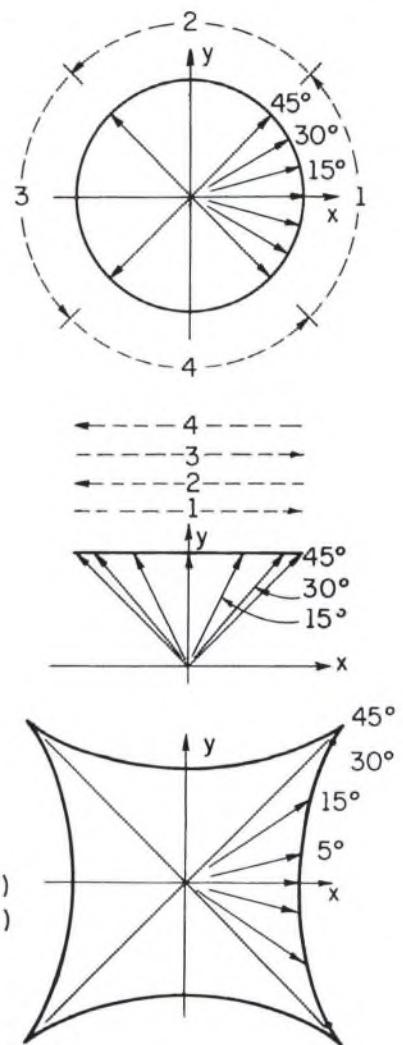


FIG. 1. Dependence of tracing distortion upon direction of modulation.

to correct their errors, and adjust their simulator in agreement with the more complete developments given elsewhere.⁶ The present author is particularly earnest in this hope, not only because he personally admires the effort these men have undertaken, but also because he has had the opportunity to hear the result, in before-and-after pressings. The corrections made, however incomplete, combined with an evidently high level of workmanship, seem to produce a remarkable improvement, especially notable with the better pickups.

⁶ D. H. Cooper, "Power Series Corrections of Tracing Distortion," *IEEE Trans. on Audio AU-13* (in press) (1965).

Continuous Delay Regulator for Controlling Recording Errors*

DUANE H. COOPER

University of Illinois, Urbana, Illinois

The principle that tracing and tracking errors arise solely from delay-modulation processes has not found wide application for controlling these errors, because of device imperfections. These may be overcome by using a regulation principle. Regulation may also be used in a switching-synthesis scheme, achieving continuous delay control without using delay lines. The regulation principle invoked is that of the phase-locked loop. The switching-synthesis scheme is a sampled-data analog of the phase-locked loop. Rational design approaches are indicated.

INTRODUCTION

DISTORTION arising from tracing and tracking errors in phonograph reproduction has been shown to arise from nothing more than a delay modulation of the signal, for which the instantaneous delay variation is directly controlled by the modulated signal itself. Further, delay modulations of just this form may be introduced in a complementary way to exactly cancel the delay variations caused by these errors.[1,2] Figure 1 shows, schematically, the style of modulation that is involved.

In essence, two independent modulators are required for the correction of tracing error, one for each of the two signals to be impressed on the two groove walls, and it is the velocity signals that are to be delay modulated and that are to control the modulations. In contrast, a synthetic adjustment of cutting angle to control tracking error requires that the two displacement signals suffer a common delay modulation, controlled by a linear combination of displacement signals that have been so modulated. The particular linear combination depends on the direction in which the angle adjustment is to be imposed; it is the difference for vertical angles.[3]

So far as the present author is aware, no actual use of the schematic procedure of Fig. 1 has ever been reported for the control of tracing error. As is well known, three other schemes have been tried, and two of these are evidently in commercial practice.[4,5] By the very nature of tracing error, these four schemes must all come to the same thing,

of course, namely delay modulation. Even so, the three other schemes do differ widely among themselves in terms of immediate principles and techniques exploited as well as in the closeness with which their immediate principles approach the fundamental process of Fig. 1.

The oldest of the three is the reverse-poled dubbing procedure of MacNair.[6] Though the scheme may be analyzed using the delay-modulation concept, the only immediate principle involved is that the tracing error generated with reversed signal polarity is complementary to that ordinarily generated, as a simple geometric argument will demonstrate.[7] The procedure has not attained widespread practice, because of the difficulties involved in high-quality disc-to-disc dubbing.[4]

The next oldest scheme, in concept, is the harmonic-

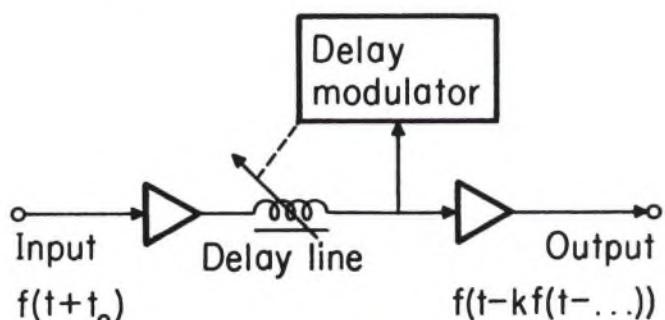


FIG. 1. Delay modulator flow chart. The scheme achieves a recursive insertion of delay variations, indicated by the expressions shown, as required for correction of tracing or tracking error. Though the topology of a feedback system is needed, no regulation is implied for the delay values themselves.

* Presented October 14, 1965 at the Seventeenth Annual Fall Convention of the Audio Engineering Society, New York.

generation simulator of Redlich and Klemp.[5] Again, it may be analyzed using delay-modulation principles, by representing the delay variation in a Taylor's-series expansion.[8] However, such expansions may be made, as Lewis and Hunt have shown, without using the modern concept of delay modulation.[9] The harmonic representation, which is the immediate principle, stems directly from those early expansions. Evidently, nonlinear devices of requisite stability have only recently become available to make the application practicable.

The most recent, again in concept, of the three, is the recording correlator of Fox and Woodward.[4] Because of its use of delay-line switching, it approaches most closely, in immediate principle, the scheme of Fig. 1. It differs in using displacement waveforms and in switching among a discrete set of variously delayed versions in accordance with an envelope-generation principle. The use of displacement waveforms rules out pure delay modulation; an amplitude-modulation term is also needed to correct the displacement error,¹ which term is correctly generated in the correlator by bias switching. It is, in fact, possible to describe the correlator principle without using the modern conception embodied in Fig. 1.[10]

While pure delay modulation has not been used for the correction of tracing error, it has quite recently been used to adjust the effective cutting angle so as to control tracking error in making a vertically modulated recording as a pickup-testing vehicle. The application is reminiscent of the first similarly limited commercial use of the harmonic-generation principle.[11] It is not known whether two-channel recordings have been made. The scheme allows for only a discrete adjustment in delay variations, providing a fading style of commutation among delay-line taps.[12] In this way, evidently, certain device difficulties associated with the scheme of Fig. 1 have been overcome.

These device difficulties have to do mainly with the variable delay line, or its equivalent. The mastery of these difficulties is a matter of engineering the devices themselves and making engineering compromises so that available devices may be exploited, as exemplified by the fading commutation. Though such engineering will always be valuable, there is an alternative, namely the design of a system whose performance is not so sensitive to the properties of the devices used.

Such systems usually employ feedback to fulfill a regulation principle. The system of Fig. 1 has the topology of a feedback system, but it does not regulate the delay, though it does achieve the recursive insertion of signal-controlled delay. The varying delay so generated is not compared to the signal values in such a way as to remove any error in the formation of the delay. Thus, any errors inserted by the delay modulator will remain uncorrected.

The next section will show how a continuous regulation of instantaneous delay variations may be achieved, using the principle of a phase-locked loop.[13] It will be seen that the regulation tends to overcome faulty behavior in the devices. Since an approach to delay overload is an important desideratum in these schemes, the mathematical

theory of overload is developed and its spectral implications are discussed in the light of requirements to be met by the phase-locked loop.

A switching-synthesis scheme is next shown as an analog of the phase-locked loop, replacing the delay line by a multi-address storage unit, in which the delay regulation is controlled by switching address locations for storage and emission with appropriate timing. The system exploits the sampling theorem,[14] as does another all-pulse scheme, the parallel-channel skew sampler. The sampling theorem is reviewed to show rational design procedures to be used in meeting performance specifications. The use of these systems in controlling recording errors is then discussed.

DELAY REGULATION IN A PHASE-LOCKED LOOP

Figure 2 shows a phase-locked loop using a phase comparator which detects a pilot-signal phase error and regulates it toward zero by supplying control signals to a vari-

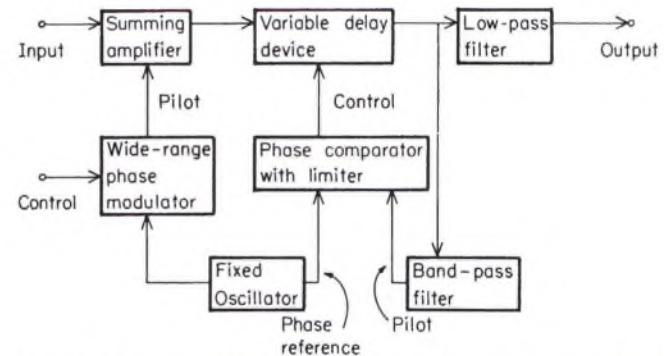


FIG. 2. Delay regulator flow chart. The phase variations produced by the control signal are nulled by regulation in a phase-locked loop, achieving phase control *via* a variable delay device. The system achieves the recursive delay insertion of Fig. 1, and at the same time removes, *via* regulation, any nonlinearities in the response to the control that may be exhibited by the variable delay device. An AGC loop may be added.

able delay device. In so doing it removes the phase modulations introduced in a pilot control signal and, in this way, regulates the delay imposed on the input signal, also transmitted through the same delay device to be presented at the output. It differs from usual loops in achieving phase control by delay adjustment rather than *via* a deviable oscillator or similar device.[13] This style of regulation achieves a recursive delay insertion, as may readily be seen. Suppose the phase modulator produces the pilot signal of nominal frequency ω_p :

$$\cos[\omega_p t + C(t)], \quad (1)$$

in which $C(t)$ is the input control signal. Then, if the phase comparator causes the delay variation to be $D(t)$, the pilot signal emerging from the variable delay device must be

$$\cos[\omega_p [t - D(t)] + C[t - D(t)]]}. \quad (2)$$

The net phase modulation, then, is

$$\phi(t) = -\omega_p D(t) + C[t - D(t)]. \quad (3)$$

However, it is just this $\phi(t)$ which is to be regulated to zero. Thus, the delay variations which have been introduced

¹ See, for example, the discussion following Eq. (2) of Ref. 1.

are described by the recursive form

$$D(t) = \tau_p C[t - \tau_p C(t - \dots)], \quad (4)$$

in which τ_p has been written for $1/\omega_p$. This is of the same form as the recursive delay insertion shown for Fig. 1. The signal transmitted along with the pilot signal will suffer the same delay variation as in Eq. (4).

A variety of variable delay devices could be used for the system of Fig. 2. Of these, the ones offering an electrical control of the delay would have the greatest appeal for an all-electronic system. An example would be the varactor delay line.[15] This happens also to be a prime example of a scheme presenting device difficulties as cited above. The line resembles a conventional L-C delay line except that the capacitance is supplied by varactor diodes with variable bias. Provisions are also sometimes made for the use of saturable inductors, with a variable bias flux controlling the inductance values, as well. As would be expected, the parameter variations are hardly linear with the electrical control values, and unless the inductance and capacitance values can be made to track each other there is also a spurious variation in characteristic impedance. Further, for low distortion, the signal level must be small compared to the variation in the control signal so that there is a tendency for the delayed signal to be contaminated by the control signal (crosstalk).

It is possible and not unduly complicated to supply the control signal in a balanced way to help defeat the crosstalk problem, but there remains the impedance variation, the nonlinearity of the control, together with yet one other problem, the duplex-phase modulation problem, to be discussed shortly. The impedance variation causes two effects. One is spurious amplitude modulation and the other is reverberation caused by transient mistermination. If the mistermination is not too severe, the line can be made sufficiently lossy to defeat reverberation. Again if the mistermination is not too severe, the spurious amplitude modulation may also be removed by incorporating an automatic gain control (AGC) loop into the system, sensing the level of the pilot signal.

The control nonlinearity may be matched against a complementary, deliberately introduced nonlinearity, leaving a residuum to be removed by the feedback inherent in the phase-locked loop. It will be seen that this feedback also tends to defeat the duplex-phase modulation problem. Duplex-phase modulation is a characteristic of variable propagation-speed delay devices. These devices may be likened to a tape-loop delay device in which the playback head is a fixed distance "downstream" from the recording head. The playback head picks up the signal just recorded, but with a delay depending on the tape speed. If the speed is varied without stretching the tape, the delay will be varied but the signal will have its time scale altered, not only at the playback head, but also in a complementary way as it is recorded at the recording head. Thus, variations in phase introduced at a given moment, to satisfy some condition at the playback head, will reappear at the same time in the future when the corresponding variations also introduced at the recording head eventually arrive at the

playback head. This is the duplex-phase modulation. Fortunately, the system of Fig. 2 keeps a detailed account of these phase variations, too, and straightens them all out at the end provided delay overload is never quite attained.

A close approach to delay overload is, of course, to be desired, since for tracing error it corresponds to the approach to curvature overload. The full significance of delay overload will be explained in the next section. Suffice it to say here that delay overload may be approached as closely as one pleases, in principle, provided the delay device can accommodate the requisite bandwidth and provided the comparator can follow the rapid phase transients involved without risk of breaking the phase lock. Detailed examination of these problems, together with the stability problems characteristic of regulator systems, is beyond the scope of this paper. Thus, while it does not appear that these problems impose requirements beyond the state of the art, it will be assumed that these requirements have been softened for the system of Fig. 2 by demanding an approach only half-way towards overload. Then, if full overload is required, two such systems may be connected in cascade.

THEORY OF DELAY OVERLOAD

Delay or phase overload begins when the total instantaneous phase of a periodic signal, such as the pilot signal of Eq. (1), namely $\psi(t) = \omega_p(t) + C(t)$, fails to be strictly monotone increasing. The condition for avoiding delay overload then is

$$\psi'(t) = \omega_p + C'(t) > 0, \quad (5)$$

where the prime is used to denote differentiation. In Eq. (5), $\psi'(t)$ is the total instantaneous radian frequency and $C'(t)$ is the instantaneous frequency deviation, in accordance with the usual identification of instantaneous phase rate with instantaneous frequency. Thus Eq. (5) represents the intuitively plausible condition that the downward frequency excursions must obey

$$-C'(t) = \Delta\omega(t) < \omega_p, \quad (6)$$

restricting frequency deviations to be smaller in peak value than the carrier frequency, so that the total instantaneous frequency shall always be positive. This same condition also prevents the intervals between the zero crossings of the oscillation of Eq. (1) from becoming unboundedly long, or requiring an interpretation of intervals as presented in a reversed time sequence.

The same restriction applies to Eq. (4); the rate of change of delay is[16]

$$D' = \tau_p C / (1 + \tau_p C'), \quad (7)$$

and this will be infinite for $\tau_p C' + 1 = 0$, i.e., for a violation of Eq. (5). An even greater violation of Eq. (5) would force $D(t)$ to become double-valued, an impossibility related to the impossibility of time reversal. It is this consideration of Eq. (7) that is the source of the term "delay overload," coined in analogy with the well-known term, "curvature overload," applied to tracing error. The expression for tracing error has the same recursive structure, generally, as Eq. (4), except that the role of C would be played

by slope values. Then the considerations involving Eq. (7) would involve slope-of-the-slope, i.e., curvatures. Examined in a little more detail, it is seen that curvature overload obtains when the peak curvature of the recorded waveform matches that of the playback stylus, so that the result of the tracing shows points of infinitely sharp curvature.

In the art of tracing correction, curvature overload describes the imposition of corrections sufficient to impose the same sharp curvatures on the maxima (i.e., projections outward from the groove wall) of the wall-modulation waveform. Upon playback, these sharply turning maxima are of course converted by the tracing error to smoothly turning maxima. A similar slope limitation may be expressed for a waveform to be recorded with an inclined cutter; however, this slope overload can never be obtained in practice because of other factors limiting slope.[16] For the shorter wavelengths, curvature overload is the only limitation in maximum curvature, so that any tracing correction to be applied to signals to be recorded at the highest levels that do not entail an information loss must be able to provide delay rates just short of overload.

For a peak phase excursion C_0 , undertaken sinusoidally, the peak frequency excursion $\Delta\omega$ is

$$\Delta\omega = \omega_0 C_0, \quad (8)$$

in radians per second for ω_0 as the modulating frequency, also in radians per second and C_0 expressed in radians. The peak delay value then is $D_0 = \tau_p C_0$, so that replacing C_0 by $D_0/\tau_p = D_0 \omega_p$, avoidance of delay overload with sinusoidal modulation imposes the condition, from Eq. (6), that

$$\omega_p \omega_0 D_0 < \omega_p,$$

or

$$\omega_0 D_0 < 1, \quad (9)$$

or, writing τ_0 as the radian interval, $\tau_0 = 1/\omega_0$,

$$D_0 < \tau_0, \quad (10)$$

a condition which makes no reference to the frequency of the phase-modulated pilot signal in specifying the maximum permissible rate, as in inequality (9), with which a given peak delay may be modulated.

Applying the same condition to delay automodulation in a sinusoid, i.e., setting $\omega_0 = \omega_p$, one sees that overload is attained for a peak phase excursion of one radian. Fourier analyses have been made for delay-automodulated sinusoids, and they show the Fourier series to begin to fail to converge at certain points, as is to be expected at overload, for those are just the points at which the waveform is verging on becoming double-valued. Actually, at the beginning of overload, the series describes the values correctly but fails to describe the slopes at such points.[1] The series invokes only harmonic frequencies of strengths given in terms of Bessel functions. The relative strengths of these harmonic components are shown plotted in Fig. 3, for various levels ranging from the beginning of overload (0 db) to 12 db below overload. The 0 db curve shows a nearly constant slope on these logarithmic scales of 26.67 db per decade (8 db per octave), but at even 1.16 db below overload the rate of decrease with increasing harmonic number soon exceeds this.

A chart like this one can be used to estimate bandwidth requirements for the delay line. It is assumed that two systems in cascade, each like that shown in Fig. 2, are to be used, so that the overall effect will be the achievement of modulation at 0 db, i.e., verging on overload. As an illustration, the consequences of placing the pilot signal frequency at 100 kc will be examined. Let it be assumed that the highest program frequency to cause overload is 10 kc. Then, it emerges from the delay line in the second system with a 10th harmonic (at 100 kc) having a relative strength of 4.2%. On the other hand, a 2 kc tone emerges with its 50th harmonic (also 100 kc) at a relative strength of 0.5%. But at overload, the 2 kc tone has 5 times the amplitude of the 10 kc tone, so that the harmonic content deriving from the 10 kc tone should be referred to the level of the 2 kc tone. Thus, the 10 kc tone contributes 0.84% and the 2 kc tone contributes 0.5% at 100 kc. Since only a few such tones would be present at near overload strength at any one time, one should estimate that the pilot signal would have to compete with a roughly one percent background, assuming it to be present at about the same level as the delay-modulated signal emerging from the delay line. This contamination would be judged satisfactory for the phase comparator, and probably would also allow the

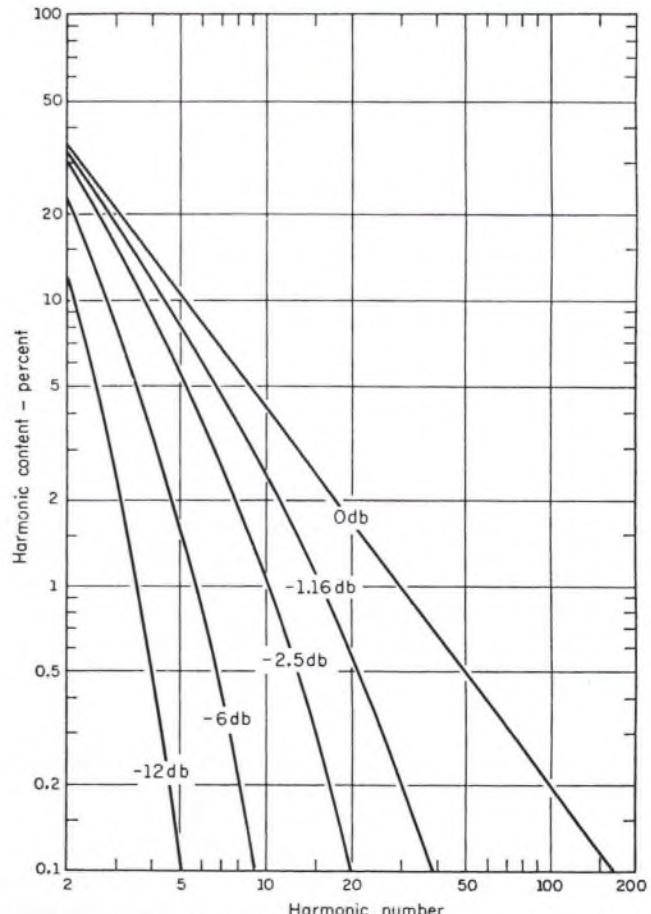


FIG. 3. Harmonic content near delay overload. The chart shows, plotted to logarithmic scales, the amplitudes of the harmonics of a delay automodulated sinusoid. The decibel figure indicates the extent of the modulation relative to the onset of delay overload. The chart is used to estimate bandwidth and sampling requirements.

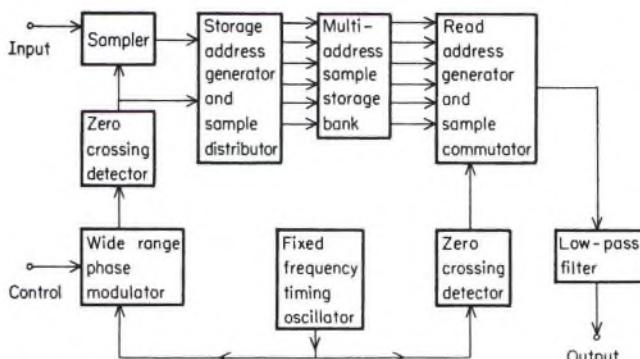


FIG. 4. Switching synthesis flow chart. This is a sampled-data analog of the regulator of Fig. 2. The zero crossings of the phase-modulated waveform are used to control the timing for the extraction and storage of signal samples into the multiaddress sample-storage bank. The zero crossings of the reference phase are used to control the emission, at regular intervals, of samples from storage according to the rule of oldest sample out first, thus delaying-out the phase modulation.

AGC loop to function satisfactorily. If the amplitude modulation to be removed were sufficiently slight, the pilot-signal level could be increased, improving the AGC action. On the other hand, the amplitude modulation sidebands would, to some degree, contaminate the emergent delay-modulated signal.

Though the pilot signal emerges free of phase-modulation sidebands because of the action of the phase-locked loop, the same is not true of the pilot signal at the input. At 6 db below overload, the peak frequency excursion is half the pilot frequency, by Eq. (6), so that a nominal delay-line bandwidth of 150 kc would be suitable. Taking 2 kc as the lowest frequency demanding overload, one calculates from expression (9) that $D_0 = 80 \mu\text{sec}$, or, at half overload, just $\pm 40 \mu\text{sec}$.² A maximum line length of four times this amount, or $160 \mu\text{sec}$, would probably be a suitable choice. A figure of merit for such lines is the product of bandwidth and line length. For the present example this figure of merit is about 24, which seems to be attainable. The conditions for the first system, in the cascaded arrangement, are not even this severe, and a willingness to restrict the overall modulation to fall 2.5 db below overload would ease the requirements still further.

The accuracy with which the delay modulation may be done is controlled by the accuracy with which the phase-modulated pilot carrier may be produced. The modulator must be also arranged so that an accurate phase reference can be produced. An equivalent frequency deviation of $\pm 50 \text{ kc}$ is not unusual, and can readily be obtained using the well-known Armstrong modulator with the normal harmonic enhancement of the modulation index, usually done in stages with appropriate limiting and filtering, combined with heterodyne adjustments of the nominal frequency.[17] Such schemes may be compensated to provide accurately linear modulation, and they readily provide for a precise reference phase to be produced.

² A more precise estimate based on a peak slope limitation of unity, set by recording stylus geometry, a playback stylus radius of curvature of 0.7 mil, and a recording radius of 2 in., gives a total delay range, peak-to-peak, of $143 \mu\text{sec}$ for tracing error correction. This is in reasonable agreement with the above numbers.

DELAY REGULATION BY SWITCHING SYNTHESIS

Although the limitations imposed by the more widely-known delay devices appear to be manageable, within the requirements of tracing error correction, it is not necessary for such limitations to restrict the application of the principle of delay regulation. In this use, the variable delay device is simply serving as a signal-storage device, of a capacity adjustable from some maximum capacity downward. A wide variety of storage devices may be used, including those used in digital computers. Since a digital representation of the signal is not a necessary element in the method of switching synthesis, however, it will be assumed that analog values are stored, since such storage is also entirely practicable.

In the use of the variable delay device, it was found necessary for the range of delay variation to span many periods of the nominal pilot signal frequency, so that the delay line would, at any instant, often be holding an extensive history of the signal being subjected to delay modulation. It will be seen that an analogous situation obtains in the switching synthesis realization. The corresponding memory system must be capable of holding a similarly extensive history of the values of the signal being subjected to delay modulation.

In Fig. 4, the memory device is represented as a Multi-Address Sample-Storage Bank. The whole system is analogous to the phase-locked loop. The new feature is the sampling operation and the storage of sample values obtained with varying spacing, to be emitted with regular spacing in accordance with the sampling theorem. The varying elapsed time between the operation of sampling and the emission of the corresponding samples from storage constitutes the varying delay. The sample values can, of course, range over a continuum, as can the delay values.

The delay values are again controlled by a phase modulator, but the phase-modulated pilot signal is now used to control the timing with which input signal values are sampled and inserted into storage. For this purpose, the zero crossings of the phase-modulated pilot oscillation serve as the occasion for generating pulses for actuating the sampler and for controlling the storage of the sample value. The zero crossings of the reference-phase oscillation are similarly used to generate pulses, timing the emergence of the samples from storage. The emergence is at regularly spaced intervals, so that pilot phase modulation has been removed, and the delay is regulated³ to conform to Eq. (4). Since the samples emerge with equal spacing, one can invoke the sampling theorem[14] to see that a suitable low-pass filter will provide a reconstruction of the continuous waveform representing the delay-modulated signal, a reconstruction that may be made as accurate as one pleases.

The address-generating scheme for distributing the samples into storage and that for commuting among storage addresses to cause the emission of the correct sample must follow the rules of the queue or waiting line: the oldest sample must be removed first. For example, the distributor inserts a new sample at the next address after the most

³ Switching synthesis achieves the regulation, however, without explicit use of feedback, thus avoiding stability problems.

recent insertion, and the commutator causes emission of a sample at the next address after the most recent emission. Always, of course, the addressing is cyclic, i.e., one always construes the next address after the very "last" to be the "first." Since the insertion frequency can never "run away," being derived from a fixed frequency *via* phase modulation, one can always choose a maximum memory capacity which will always satisfy the demand capacity, and always be assured that the memory will never be empty, except at the start. Standard initial addresses can always be set as a starting condition.

Fulfillment of the rules of the queue can be obtained in a variety of ways. One can, for example, arrange the memory addresses as in a shift register, in which the contents of one address can be propagated only to the next-most address. Then, insertion is always done at one address (the input end of the shift register) and emission from a certain other one (the other end). Special rules must govern the propagation of empty states, however. One rule might be that the "customers" waiting in the queue never allow empty spaces to remain between themselves and the emission address.

PARALLEL CHANNEL SKEW SAMPLER

The elaborate switching involved in distributing the samples among memory addresses and commutating them out again can be simplified into one operation of distributing timing pulses among parallel channels, in which each channel processes only one sample at a time. Each channel accepts the timing pulse distributed to it; produces a delayed version, the delay obeying the control signal; the delayed pulse is used to sample the input signal, which sample is stored until the next distributed timing pulse

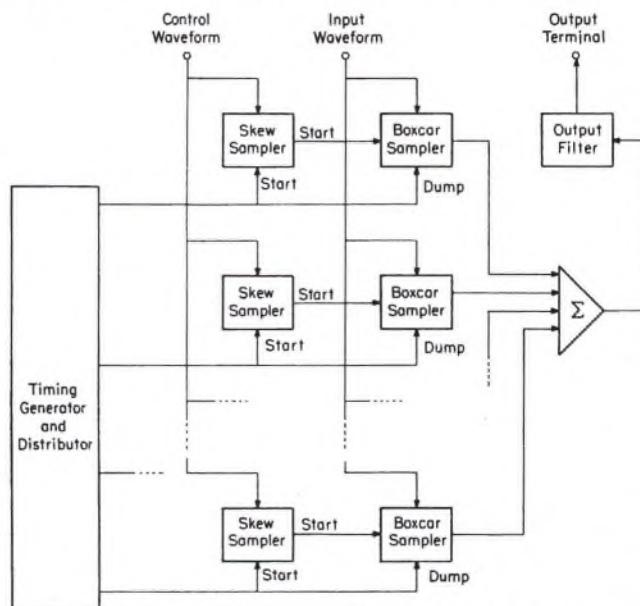


Fig. 5. Skew sampler flow chart. The phase modulator of Fig. 4 is replaced by a serrasoid pulse-position modulator, here called a skew sampler, controlling the timing for the insertion of signal samples into single storage units, here called boxcar samplers. The staggered distribution of basic reference timing pulses meets the demand for a peak delay in excess of the net sampling interval.

occurs; at that time, the sample is emitted to a common bus conducting it to the output filter, and the process begins anew.

The arrangement is shown in Fig. 5, in which each channel consists of two units, one called the Skew Sampler and one called the Boxcar Sampler.[18] The skew sampler is a pulse-position modulator, in which a timing pulse begins the generation of a linearly increasing ramp signal; at the instant the ramp value agrees with the control value, the

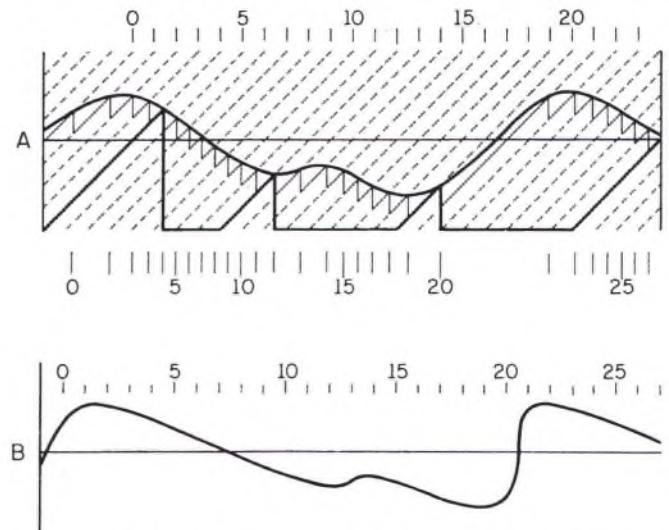


FIG. 6. Skew sampler waveforms. A. Ramp waveform generated in a skew sampler starting at a timing pulse; the second ramp shown, for example, starts at the 4th timing pulse, as numbered at the top. This is the time at which the first sample value, obtained when the first ramp achieves a value equal to the instantaneous control waveform value, is scheduled to be emitted from the boxcar sampler. The sample was obtained, however, at the time of the 4th position-modulated pulse, as numbered at the bottom. The portions of ramps shown in lightweight lines, aligned with the evenly-spaced sloping dashed lines, indicate, in part, the action of the other seven (for example) skew samplers. B. Plot of the sample values from A, but at equal intervals, and interpolation of a smooth curve. This shows the effect of using an input waveworm identical to the control waveform.

ramp is reset to its starting value, and a sampling pulse is emitted to the boxcar sampler. The skew sampler then awaits the next distributed timing pulse to begin its cycle anew. This style of pulse-position modulation is an adaptation of serrasoid modulation with interlace,[19] except that the merging of the interlaced samples is not done until after boxcar sampling.

Meanwhile, the boxcar sampler measures the value of the input waveform at the instant it receives the sampling pulse from the skew sampler, and holds this measured value, or waveform sample, in a storage unit, such as a capacitor. Upon receiving the next distributed timing pulse, it emits the sample to the common bus conducting the sample to the output filter. It then awaits the next sampling pulse from the skew sampler.

The Timing Generator and Distributor generates timing pulses at regular intervals and distributes these cyclically and in sequence among the n channels so that each receives its distributed timing pulse also at regulated intervals, which intervals are n times as long as that generated. The com-

mon bus merges the emitted samples for conduction to the output filter, which then receives samples at a fixed rate, that of the timing generator.

The operation of this system is illustrated in Fig. 6. Figure 6A shows a control waveform with the ramp waveform of one of the skew samplers shown in bold lines. The lighter lines show portions of the ramps of other skew samplers in the system. Only portions are shown, to avoid showing a multiplicity of overlapping waveforms; however, dashed sloping lines representing extensions of the ramp portions have been drawn to show the regularly repeating basis for the ramps and to show the bounds on delay and, hence, on control values that may be accommodated. The similarity of these sloping lines to those of a skew grid is responsible for the name, skew sampler. The times of occurrence of the timing pulses are shown at the top, and the times of production of the position-modulated sampling pulses are shown at the bottom.

The skew sampler chosen for illustration began the first ramp, shown, four timing pulses prior to the one marked "0," and is scheduled to begin the next one at the timing pulse marked "4." Since the latter marks the time at which the sampled value is scheduled to emerge from storage, the sampling pulse generated when the ramp value agrees with the control value (the control value occurring at the same instant) is also marked "4," in the position-modulated train. This same skew sampler is also responsible for the production of the 12th, 20th, and 28th (the last not shown) position-modulated sampling pulses. The case illustrated is for eight parallel channels.

For purposes of illustration, it was assumed that the waveform supplied for boxcar sampling had the same form as that controlling the skew sampler. Since these samples emerge from the boxcar samplers at equal intervals, Fig. 6B was constructed by plotting these sample values, read from Fig. 6A at the sampling points but plotted at equal intervals, and placing the 0th one to suit convenience. The action of the output filter was then simulated by interpolating a smooth curve among these points. The result is, of course, a strongly skewed version of the curve in Fig. 6A. The region between samples 20 and 21 marks a close approach to overload, and demands of the output filter a transition in values that is nearly as rapid as it can supply.

It is evident that each channel uses elements well-known in the art of pulse circuitry,[20] and that these functions may be rather simply and accurately realized. Indeed, one may readily imagine a modular construction using low-cost integrated circuits, so that a large number of channels may be economically and compactly accommodated in one unit. The number of channels required for a given application may be determined from the theory of sampling.

SAMPLING REQUIREMENTS

The sampling theorem which these switching synthesis schemes must obey is also the basis for the familiar pulse-code modulation systems.[14] It will be necessary to review it only briefly here, in preparation for a discussion of the problem of controlling the intensity with which "alias" frequencies will be generated in the delay-modulated signal.

Figure 7 shows, schematically, certain waveforms, together with their corresponding frequency plots or spectra, as occur in sampling or pulse amplitude modulated systems. In Fig. 7A there is shown an arbitrary waveform and a schematic representation of its spectrum, the latter on the left. If the waveform were periodic, the spectrum would consist only of discrete frequencies. Otherwise it would be continuous; the schematic representation is designed to be ambiguous on this point, since the sampling theorem is indifferent also. Let it be supposed initially, however, that there is a certain maximum frequency, $f_m = f_{max}$, characteristic of this waveform, i.e., that it be band limited.

Part B of Fig. 7 shows a periodic pulse train of frequency $f_{pulse} = f_p$ on the right, and a schematic representation of its frequency analysis on the left. This time, there certainly are only discrete frequencies. There are a dc component, a fundamental component at f_p , a second harmonic at $2f_p$,

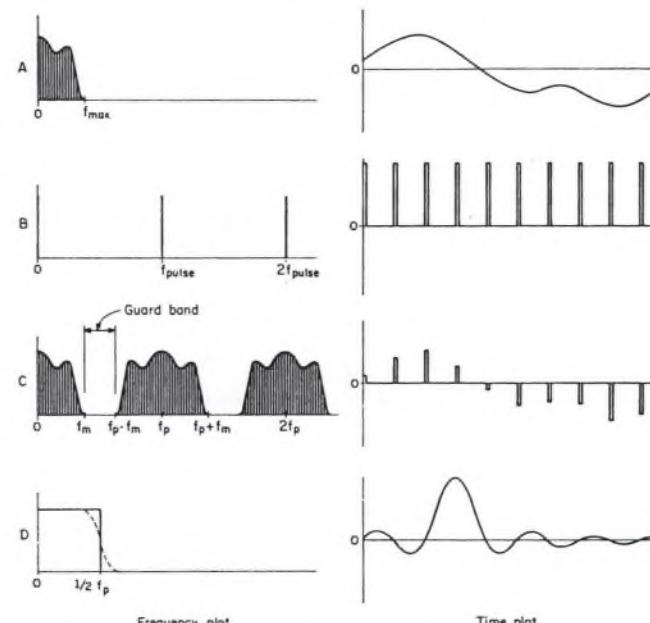


FIG. 7. Sampling waveforms with frequency plots. A. A waveform together with its schematic frequency plot. B. A pulse train together with its frequency plot. C. The result of amplitude modulating the pulse train shown in Fig. 7B by the waveform of 7A. It is seen that each of the frequency components of the pulse train blossoms out into the double-sideband spectrum characteristic of amplitude-modulated carriers. D. The transient response and corresponding frequency response of a sharp-cutoff low-pass filter. It is seen that the transmission of the waveform of Fig. 7C through such a filter will restore the frequency plot of Fig. 7A and thus reconstruct the waveform exactly from its samples.

etc. Figure 7C shows the result of amplitude modulating the pulse train of Fig. 7B by the waveform of 7A. It is as if each of the frequency components of the train were a carrier subjected to the same amplitude modulation, so that the characteristic double-sideband-modulation spectrum now appears centered at each component frequency of the original pulse train. The modulation of the wavetrain itself is depicted on the right. The height of each pulse represents a value sampled from the waveform of Fig. 7A.

It is assumed that f_p is more than twice f_m , so that a clear gap, called a guard band, appears in the spectrum of

the sampled waveform of Fig. 7C. This gap makes it possible to use a low-pass filter cutting off at f_m , or more generously, at $\frac{1}{2}f_p$, as on the left in Fig. 7D, to recover the original spectrum of 7A and thus reconstruct the waveform of 7A exactly from its samples. The characteristic $(\sin x)/x$ transient response of such a filter is plotted on the right. What happens to the pulses is that the transient response is used as an interpolation or weighting function, in a weighted moving average (convolution) over the sample values, to construct a smooth curve exactly reconstructing the waveform of Fig. 7A. Since a filter which cuts off so sharply is not physically possible, a guardband of finite width would be needed to allow the use of a filter showing a more realistic cutoff, as shown by the dashed curve.

What makes the sampling theorem possible is the fact that the superposition of a maximum frequency constrains the extent to which the waveform of Fig. 7A may be supposed to be free to "wiggle around" between samples, so that the filter "knows," from the information in its transient response and in the sample values, just how to reconstruct the in-between values. It is necessary, since the filter has a fixed transient response, that the samples be presented to it with unvarying spacing. Analysis corresponding to the following would be very difficult otherwise.[10]

Unfortunately, real-life waveforms do not have a definite maximum frequency, and no filter can make them so, since no filter can be made that has infinite attenuation everywhere outside a given band. There will be "tails" to these spectra, extending from the intended bands into the bands in which the appearance is not wanted. Thus, in Fig. 7C, the band to be selected will contain spurious or "alias" frequencies resulting principally from spectrum "tails" extending down from the band centered at f_p , and no amount of filtering will remove these spurious components. What is possible is to choose the filter characteristics and pulse frequency in a coordinated way, to make the alias level as low as will satisfy some criterion of tolerance for interpolation error.[21].

This is the case, of course, with the samples being presented to the output low-pass filter of Figs. 4 and 5. The timing waveform has been used as reference to delay-out the pulse-position modulation, so that these pulses will be presented at equal intervals, but these pulses now are samples representative of a delay-modulated waveform. No matter how carefully the original waveform might have been filtered, its spectrum would have been widened by the delay modulation, so that it is to the widened spectrum that the alias frequency considerations, together with those of Fig. 7, must pertain.

So long as the modulating or control waveform has a spectrum of no greater nominal width than that of the original input waveform, the widening of the spectrum is no greater than that due to delay automodulation, which, in turn, is well characterized by the generation of harmonics in the delay automodulation of a sinusoid, as plotted in Fig. 3. Further, there will be only a few frequencies present at near overload strength at any one time, so that Fig. 3 may be used to estimate the strengths of the alias components.

As an example of these alias considerations, consider a sampling frequency $f_p = 140$ kc. At the beginning of overload, it is seen that the 9th harmonic of a 15 kc tone, i.e., 135 kc, will be present with a relative strength of 4.8%. As a lower modulation side tone on the

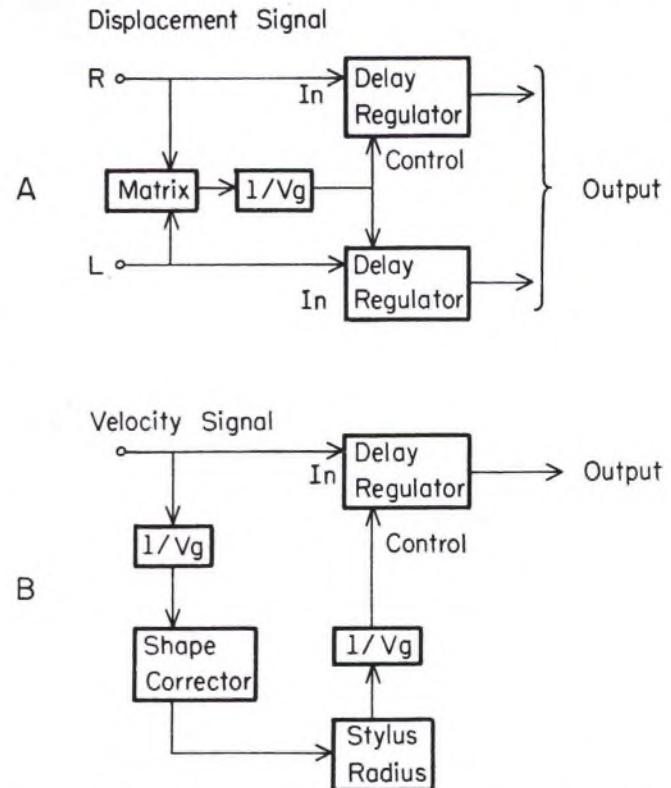


FIG. 8. Flow charts for controlling recording errors. A. Procedure for synthetic adjustment of cutting angles for the control of tracking error. The control gain varies as $1/v_g$, but may also be adjusted to control the exact amount of angle correction. The matrix box may be adjusted to obtain the correct balance between lateral and vertical angle adjustment. B. One channel of the system for controlling tracking error. The shape corrector supplies a slight memoryless non-linearity to allow for departures in stylus profile shape from the linear slope function characteristic of a parabolic profile. A duplicate system for the other stereo channel is needed.

*140 kc carrier, this would appear at 5 kc. Relative to a 15 kc tone, a 5 kc tone near overload would have triple the strength of the 15 kc tone, so that the relative strength should be taken to be 1.6% for the alias. At 2.5 db below overload, the corresponding figure would be 0.3%. On the other hand, a 5 kc tone verging on overload produces its 5 kc alias by way of its 27th harmonic for which the strength would be 1.1%, but at 2.5 db below overload it would be about 0.03%. Thus, it would appear that 140 kc is a reasonable choice for the sampling frequency.

It is clear that the exact choice of sampling frequency is to be specified in terms of a specified error criterion for interpolation, via a tolerance level for the alias components, together with a specification of the degree to which overload may be approached. For example, if one were content to approach only to within 12 db of overload, one could satisfy a 0.1% tolerance with a sampling frequency of only 63 kc,

but the same tolerance at 0 db would require a sampling rate of 798 kc. On the other hand, a choice of 49 kc would be consistent with a 1% tolerance at 12 db below overload. It is a special merit of these sampling schemes that they lend themselves readily to a rational prediction of performance accuracy for given choice of design parameters.

These examples of sampling frequency have been quoted as multiples of 7 kc, because this multiple describes the number of memory cells needed for sample storage in the system of Fig. 4, or the number of parallel channels for Fig. 5, if one assumes a peak-to-peak delay range of 143 microseconds, the period corresponding to 7 kc. Thus, for the 140 kc sampling rate, one is anticipating a peak-to-peak delay range of at most 20 sample periods, so that 20 storage addresses must be available, or 20 skew-sampling channels. This number, then, is directly related to memory cost.

CONTROL OF RECORDING ERRORS

As is evident, these delay regulating schemes may be used to control effective cutting angle to control tracking error, in addition to the use of controlling tracing error. The angle control may be used to radically alter the effective cutting angle natural to the cutter, without altering the mounting of the cutter or altering in any way its normal dynamic cutting characteristics. Alternatively, if the natural angle is nearly the desired one, the angle control may be used to trim the angle to more exact conformance with a given standard. It may happen, for example, that the angle is nearly correct vertically but shows some variation with cutting radius[3] or with depth of cut,[22] or shows a small lateral component.[5] These variables all may be brought under control automatically, if desired.

Angle control does require two independent memory systems, but only a common control.[3] Thus, in the system of Fig. 5, the skew sampler may be common to both stereo channels. This makes for some economy, as would also be the case if only modest angle adjustments were required, since the memory cost would be proportional to the range of angle adjustment. Given sufficient memory, an adjustment of gain for the control channel selects the angle at a given groove speed. For a constant angular adjustment, of course, the gain must be proportional to the reciprocal of the groove speed. (See Fig. 8A.)

Consider a 2.4° adjustment. The maximum delay modulation occurs for the lowest frequency which may be recorded at maximum slope. Let this be 500 cps, for which the radian interval is $\tau_0 = \pm 317 \mu\text{sec}$. This would be the delay required at overload, which would come for a combined condition of unit slope and a 45° adjustment. Unit peak slope is desired, but the adjustment is only 2.4°, for which the tangent is 0.042. Thus, the maximum delay demand is for $\pm 13 \mu\text{sec}$, a peak-to-peak range of 26 μsec . The condition is 28 db below overload, so that aliasing from this source is not a consideration. One may sample at 38 kc, for example, the frequency corresponding to a 26 μsec period, and use only one storage device. (A 23° adjustment would require 10 storage devices and would require an approach to within 7.5 db of overload, for frequencies as high as 5 kc. However, the 7th harmonic of 5 kc is at

35 kc and has a relative strength of about 0.2%, so that the sampling frequency probably would not have to be reconsidered.)

Delay regulation for the control of tracing error does require independent delay control of each of the two stereo signals, though the two regulators may share a common source of timing. In the skew sampling scheme, the timing distributor may also be shared. The control signal is proportional to velocity if the simulated stylus is parabolic. Otherwise, a slight memoryless nonlinearity must be introduced to express the slope function of the simulated profile.[8] This simulation must take into account the effective cutting angle of the system to follow tracing correction.[16] For this reason, it is simpler for the angle control to precede tracing correction.[1] The input to the shape correction nonlinearity should be the recording velocity scaled to (divided by) the groove speed, to simulate slope. The output from the shape corrector should be proportional to the radius of curvature of the simulated stylus and again scaled to (divided by) the groove speed, to simulate a delay to be inserted into the control channel. The shape corrector may use a simple diode-resistor network. A possible refinement would be the allowance for the finite tip shape of the cutting stylus in this shape corrector. In any case, the system layout of Fig. 8B shows the functions described for a single channel.

There are, of course, certain precautions to be observed. Any fixed delays that will occur must be identical for the two channels. For example, the low-pass filters will involve such a delay; the filters should be identical between the two channels. Any other filtering, such as to achieve a certain recording characteristic, should precede the whole correction system. The same is not the case for equalizers designed to provide a flat-amplitude and linear-phase frequency response for the cutter. These must be considered a part of the cutter, and should be adjusted to achieve these characteristics over as wide a frequency band as possible. Such cutter equalization is a more crucial matter when proper tracing correction is to be undertaken, since the correction may be viewed as having performed a nice adjustment of Fourier components over the whole of the audio band in just such a way as to cancel the components generated by the playback errors. There are also a whole host of workmanship factors such as control and frequent inspection of the cutting stylus shape, which are just as crucial but which are beyond the scope of this discussion.

It may be concluded that these delay-regulating schemes offer rich opportunities for hardware realizations. They offer rational bases for design to achieve precise, stable, continuous control of instantaneous delay. The vagaries of existing relay lines may receive automatic compensation through regulation, or they may be dispensed with entirely, and reliable pulse techniques substituted, in the switching synthesis and sampling schemes. The prospect is for reliable and rigorous control of disc recording errors to accommodate high-peak recording levels at sharply reduced distortion, so that one may exploit to the full the means now becoming available for producing low-distortion master tape recordings of extended dynamic range.[23,24]

REFERENCES

1. D. H. Cooper, "Unified Analysis of Tracing and Tracking Error," *IEEE Trans. on Audio AU-11*, 141 (1963).
2. D. H. Cooper, "Integrated Treatment of Tracing and Tracking Error," *J. Audio Eng. Soc.* 12, 2 (1964).
3. D. H. Cooper, "On Tracking and Tracing Error Measurements," *J. Audio Eng. Soc.* 12, 312 (1964).
4. E. C. Fox and J. G. Woodward, "Tracing Distortion—Its Cause and Correction in Stereodisk Recording Systems," *J. Audio Eng. Soc.* 11, 294 (1963).
5. H. Redlich and H. J. Klemp, "A New Method of Disc Recording for Reproduction with Reduced Distortion: the Tracing Simulator," *J. Audio Eng. Soc.* 13, 111 (April, 1965).
6. MacNair, see discussion, *J. Soc. Motion Picture Engrs.* 21, 182 (1938); follows J. A. Pierce and H. V. Hunt, "Distortion in Sound Reproduction from Phonograph Records," *loc. cit.*, pp. 157-182.
7. D. H. Cooper, "Compensation for Tracing and Tracking Error," *J. Audio Eng. Soc.* 11, 318 (1963).
8. D. H. Cooper, "Power Series Corrections of Tracing Distortion," *IEEE Trans. on Audio AU-13*, 27 (March-April, 1965).
9. W. D. Lewis and F. V. Hunt, "A Theory of Tracing Distortion in Sound Reproduction from Phonograph Records," *J. Acoust. Soc. Am.* 12, 348 (1941).
10. D. H. Cooper, "Construction of Correlator Waveforms," *J. Audio Eng. Soc.* 13, 236 (July, 1965).
11. B. B. Bauer and A. Schwartz, "A Record for Measuring Vertical Pickup Angles," presented October 17, 1963 at the Fifteenth Annual Fall Convention of the Audio Engineering Society, New York.
12. E. C. Fox, "An Electronic Delay Modulator," presented April 28, 1965 at the Twelfth Annual Spring Convention of the Audio Engineering Society, Los Angeles, California.
13. T. J. Rey, "Automatic Phase Control: Theory and Design," *Proc. IRE* 48, 1760 (1960). See also 49, 590 (1961) and 53, 494 (Letter, May, 1965).
14. B. M. Oliver, J. R. Pierce, and C. E. Shannon, "The Philosophy of PCM," *Proc. IRE* 36, 1324 (1948).
15. See, for example, variable delay lines offered by Andersen Laboratories, 501 New Park Avenue, West Hartford, Connecticut.
16. D. H. Cooper, "Interaction of Tracing and Tracking Error," *J. Audio Eng. Soc.* 13, 138 (April, 1965).
17. A. Hund, *Frequency Modulation* (McGraw-Hill Book Company, Inc., New York, 1942).
18. P. E. Harris and B. E. Simmons, "DC Accuracy in a Fast Boxcar Circuit via a Comparator," *IEEE Trans. on Electronic Computers EC-13*, 285 (1964).
19. J. R. Day, "Serrasoid Modulator," *Electronics* 21, 72 (Oct., 1948).
20. J. Millman and H. Taub, *Pulse and Digital Circuits* (McGraw-Hill Book Company, Inc., New York, 1956).
21. J. J. Downing, *Modulation Systems and Noise* (Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1964), pp. 141-149.
22. J. G. Woodward, "Reducing Distortion in Stereo Phonograph Systems," *Audio* 48, 23 (Feb., 1964).
23. J. T. Mullin, "Advanced Tape Mastering System: Electronic Features," *J. Audio Eng. Soc.* 12, 298 (1964).
24. G. S. Bahrs and W. T. Selsted, "A New Compatible Mastering Tape Recorder," presented October 15, 1965 at the Seventeenth Annual Fall Convention of the Audio Engineering Society, New York.

THE AUTHOR



Duane H. Cooper was born in Gibson City, Illinois, on August 21, 1923. He received the B.S. degree in 1950 and the Ph.D. degree in 1955, both from the California Institute of Technology, Pasadena.

Since 1954 he has been with the Coordinated Science Laboratory (then the Control Systems Laboratory) of the University of Illinois, Urbana, where he is Research Associate Professor of Physics and Electrical Engineering. In recent years his principal interests have been statistical detection, information and decision theory, with applications to pattern recognition, and the applications of digital computing systems in instrumentation and simulation of physical systems by Monte Carlo methods.

Dr. Cooper is a member of the American Physical Society and of the Audio Engineering Society.

Factors Affecting the Stylus/Groove Relationship in Phonograph Playback Systems *

C. R. BASTIAANS

Westinghouse Research Laboratories, Pittsburgh, Pennsylvania

It is shown that a phonograph pickup stylus riding in the groove of a record partly penetrates the groove walls because of elastic and plastic deformation of the record material. At high bearing loads complete plastic flow sets in and the needle leaves a permanent indentation track, while at lower loads the elastic deformation is predominant. This leads to amplitude distortion in the reproduced signal which may be of two types: one which is a function of the recorded wavelength (G function or translation loss), the other a function of the dynamic moving mass of the stylus/armature (H function or stylus/groove resonance). A third phenomenon (S function or scanning loss) is caused by the finite size of the stylus/groove-wall contact surface.

Experiments with specially built pickups show the evolved theory to be valid even for very high frequencies. Special test records with recorded frequencies up to 100,000 Hz were used for these experiments.

INTRODUCTION In this paper, the mechanical behavior of a pickup stylus[†] tracking a phonograph record groove is analyzed, using the Hertzian theories on the elastic deformation of two curved bodies in mutual contact under the influence of a force. The validity of the theories is then checked by means of measured response curves of various pickups on specially made calibrated test records.

This analysis is based on the assumptions that: *a*) the pickup is an ideal mechanical system without any spurious resonances; *b*) stylus/groove contact is always maintained; *c*) the playback stylus is a cone with a spherical tip, made of a very hard material such as sapphire or diamond; *d*) the groove profile is always symmetric with respect to the normal to the record surface; *e*) the included bottom angle of the groove equals 90°; *f*) the groove modulation is sinusoidal, in the lateral plane; *g*) the radius of curvature of the modulated groove in the lateral plane is never less than 1.5 times the stylus tip radius; *h*) nonlinear distortion components and higher harmonics are neglected; *i*) the mechanical behavior of the record material obeys Hooke's Law; *j*) the record material is homogeneous; *k*) the turntable rotates with a constant angular velocity; *l*) extraneous forces other than the vertical bearing force exerted by the stylus in

the groove are neglected (forces such as those due to stylus/groove friction, tone arm bearing friction, pinch effect, side-thrusts due to offset tone arms, non-level turntables, warped or eccentric records).

STYLUS/GROOVE FIT

Consider Fig. 1 which depicts a cross-section in a plane through the center of the stylus tip (pictured as a sphere), perpendicular to the record surface and the direction of the groove. This groove does not carry any information, i.e., it is a silent groove. Under the influence of the vertical needle force F_v , the groove walls will "give" and the needle tip will penetrate the walls until a static balance of forces is reached.

Hertz¹ has derived formulas to calculate the elastic deformation of two curved bodies in mutual contact under the influence of a force. In the case of a stylus in a groove (see Fig. 1), the equation expressing the normal force in terms of the vertical stylus force F_v reads*:

$$\delta_0 = \frac{\psi}{(2k^2)^{1/2}} \left[\left(\frac{F_v}{2\sin\beta} \right)^2 \cdot \frac{1}{R} \left(1 + \frac{R}{2\rho} \right) \right]^{1/2}. \quad (1)$$

Since $\beta = 90^\circ/2$, Eq. (1) simplifies to:

$$\delta_0 = \frac{\psi}{(2k^2)^{1/2}} \left[\frac{F_v^2}{2R} \left(1 + \frac{R}{2\rho} \right) \right]^{1/2}. \quad (2)$$

From published tables¹ we find that ψ is practically constant and equal to 2 for the geometrical configurations prevailing in the stylus/groove fit (sphere/cylinder or

* Presented October 11, 1966 at the 31st Convention of the Audio Engineering Society, New York.

† To conform with current usage, the term "stylus" rather than "needle" is used throughout this paper, despite the author's feeling that a semantic differentiation would be advisable between a (cutting) stylus and a (playback) needle—a differentiation that does exist in the other major languages.

sphere/plane). provided the radius of groove curvature is not less than 1.5 times the stylus tip radius R . Since

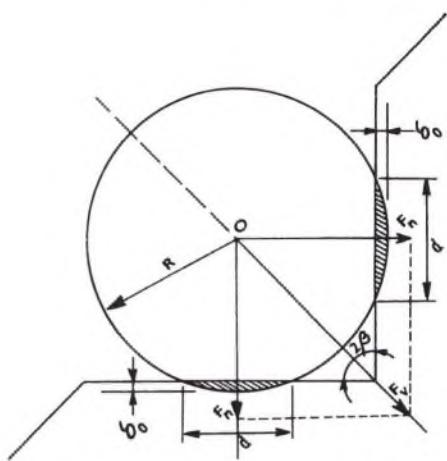


Fig. 1. Vertical cross-section of a spherical stylus tip with radius R penetrating the walls of a silent groove to a depth of δ_0 under the influence of stylus force F_r .

the stylus tip is very much harder than the record material, Hertz's coefficient k is given by

$$k = 8E/[3(1-v^2)] = (8/3)E_0. \quad (3)$$

Therefore, Eq. (2) becomes:

$$\begin{aligned} \delta_0 &= \left(\frac{9}{16E_0^2} \right)^{\frac{1}{3}} \left[\frac{F_r^2}{2R} \left(1 + \frac{R}{2\rho} \right) \right]^{\frac{1}{3}} \quad (4) \\ &= K \left[\frac{F_r^2}{2R} \left(1 + \frac{R}{2\rho} \right) \right]^{\frac{1}{3}}, \end{aligned}$$

defining K as: $K = \psi/(2k^2)^{\frac{1}{3}} = (9/16E_0^2)^{\frac{1}{3}}$.

In a silent groove, the radius of groove curvature $\rho = \infty$, so that Eq. (4) becomes:

$$\delta_0 = K(F_r^2/2R)^{\frac{1}{3}}. \quad (5)$$

The span of the contact surface or indentation width d is, according to Hertz, equal to:

$$d = 2(\delta_0 R)^{\frac{1}{2}}. \quad (6)$$

As an example, the actual value of indentation has been calculated in a concrete case. Given a pickup with the following characteristics: $F_r = 2 \times 10^{-2}$ N (approx. 2 g); $R = 17.8 \times 10^{-6}$ m (approx. 0.7×10^{-3} in.); $E = 3.3 \times 10^9$ N/m² for Vinylite (a copolymer of vinyl chloride acetate); $v = 0.35$ for most plastics; $E_0 = 3.76 \times 10^9$ N/m² (5.45×10^5 psi); $K = 3.42 \times 10^{-7}$ m^{4/3} · N^{-2/3}. Then by Eq. (5) the depth of penetration $\delta_0 = 0.765 \times 10^{-6}$ m (0.03×10^{-3} in.), and by Eq. (6) the width of contact surface $d = 7.37 \times 10^{-6}$ m (0.29×10^{-3} in.).

It should be pointed out again that the very often used expression "stylus pressure" actually refers to the stylus force. Pressure is defined as force per unit area, and a little calculation will show that the actual stylus pressure exerted by this 2 gram pickup amounts to 331×10^9

N/m² (4.8×10^4 psi). This pressure is much higher than the yield point of record plastics (around 10^8 N/m² = $14,500$ psi). It may therefore be expected that plastic flow will set in, either wholly or partly.

Figure 2 shows two straight lines giving the theoretical indentation width d caused by 2 different stylus tip radii ($17.8 \mu = 0.7$ mil and $11 \mu = 0.43$ mil) exerting normal forces in the range from .001 to .1 N (approximately .1 to 10 g) as calculated by Eqs. (4), (5) and (6).

Practical measurements of the widths of indentation tracks left on a blank Vinylite record by styli with the above-mentioned tip radii under a range of bearing loads of 1 to 10 g show that, within the range of interest, high bearing loads leave track widths that measure as predicted by the Hertzian formulas for elastic deformation (refer to Fig. 2). This leads to the conclusion that under these circumstances complete plastic flow set in after the initial elastic deformation, and a permanent indentation of a width equal to that of the elastic deformation was left.

In the lower range of bearing loads (see Fig. 2), however, the width of the (permanent) indentation track left by the stylus is much less than predicted by the theory. Viscoelastic materials like Vinylite do not exhibit a sharp division between the elastic and plastic region. Any stress will produce a response with both elastic deformation and viscous flow, the magnitudes of which depend not only on the properties of the material but also on the length of time during which the stress is applied.

A convenient way to find the track width due to both elastic and plastic deformation is to deposit a soap film on the record surface. Such a film is microscopically thin, and should not interfere with the phenomena to be investigated since it is very much softer than the Vinylite.

The needle will now leave a track with a width equal to the total deformation, purely elastic in the low range, plastic in the high range and both elastic and plastic in between. The findings are shown in the measured curves of Fig. 2. A surprising fact shows up: below a certain load the track width is even smaller than the theoretical value! Above this load, the same widths as without the soap film are found, confirming that these

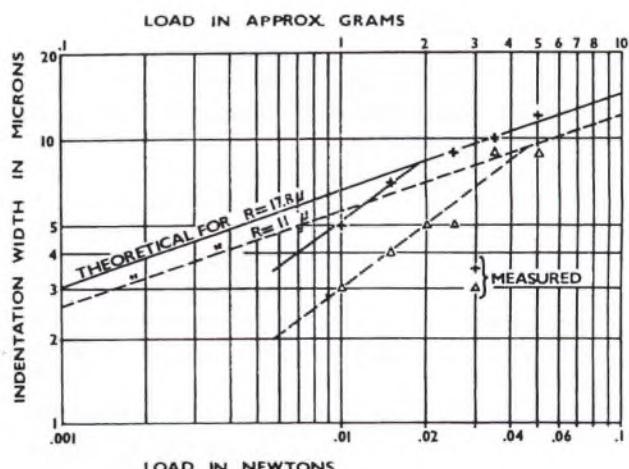


Fig. 2. Track widths left by two different stylus tips on Vinylite, calculated and measured for a range of stylus forces.

* Refer to Appendix I for definitions of all symbols.

tracks are indeed caused by complete plastic flow with very little or no elastic springback. Below the critical load the material seems to behave as a much stiffer material; note the steeper slope of the lines obtained when the measured values are connected.

This "knee" in the elastic behavior of Vinylite has been found by others, via other approaches. Hunt² has suggested the existence of a *size effect* similar to the effect observed in tiny metal whiskers that are known to be extremely strong only because of their complete lack of flaws. Barlow³ has objected to this size effect and explains the thinner permanent track widths at lower loads by elastic component recovery. His theory, however, is contradicted by our findings with the soap film. An interesting explanation was given by Walton⁴ who suggested a phenomenon very aptly termed "surfboard action." When the groove speed exceeds some critical value, the applied load apparently is supported only partially by the elastic reaction of the material, the balance being supported by the dynamic lifting action on the stylus.

This investigation has been far from exhaustive and a vast field of exploration in the rheological behavior of the record material remains. In measuring the indentation track widths on the record surface it is found that there is a variation in the value of the stiffness with position around the record and with radial distance from the center of the record. Work hardening and/or skin effect are possible causes of the anomalies sometimes found in practical measurements of indentation tracks on records.

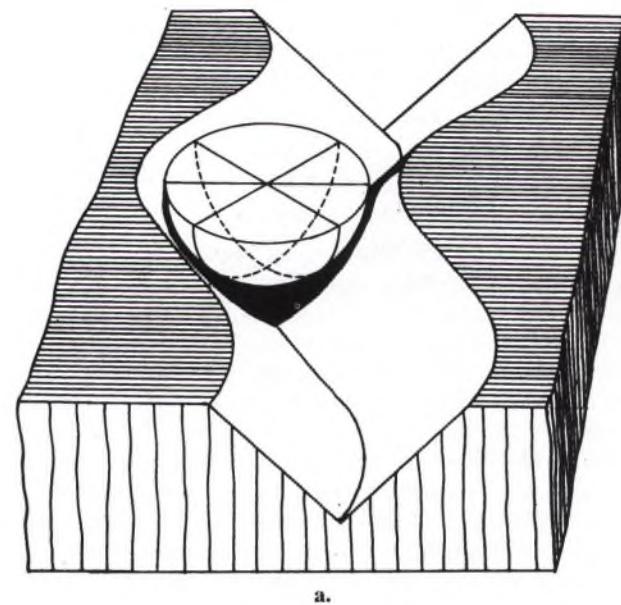
A last word on plastic flow: It is a known fact that the highest shear stress in the material below an indenter occurs at a point about half a contact-circle radius *below* the center of contact (Davies⁵). It has been suggested⁶ that in the stylus/groove contact, sub-surface yielding begins near a load of .150 g and plastic yielding at the surface starts at loads in the range from 1 to 1.6 g (using a stylus tip radius of 17.8 μ).

The fact that a phonograph record is not completely ruined by a first playing at even 5 g bearing weight is most probably due to the fact that fresh material is constantly presented to the stylus tip when the pickup plays the moving groove. Microscopic examination of a record groove that has been played once with a stylus force of only 2 g will, however, show that a slight permanent indentation track is already left on both groove walls. It has also been shown that repeated playings (with no intervening periods of time) of a small band of the record bring about an accelerated wear of the record groove. Apparently the sub-surface shear stresses are not immediately dissipated and the accumulating abuse leads to an early breakdown of the record material. Clearly the time factor plays an important part in the rheological behavior of the plastic.

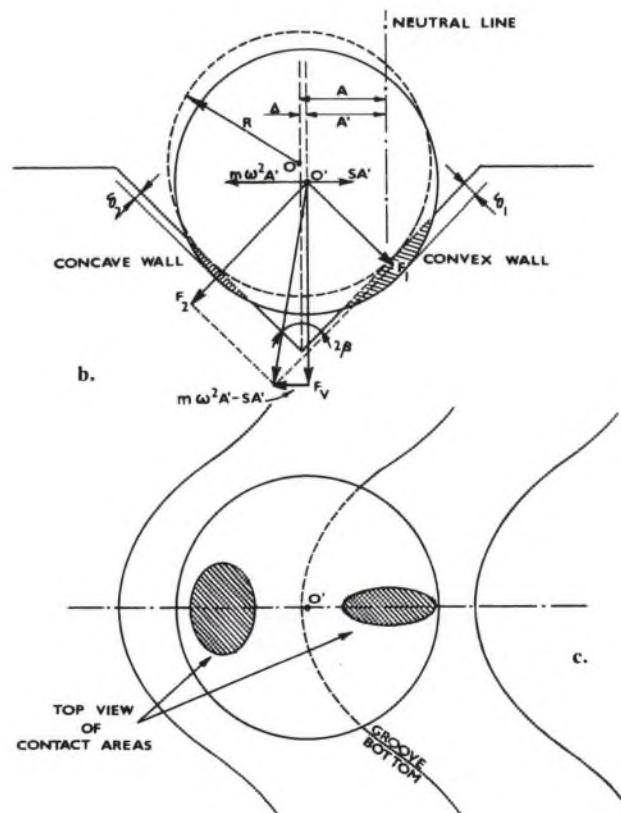
PLAYBACK LOSS

It is interesting to examine the mechanical phenomena that occur in a modulated groove because of the effects discussed above, limiting the discussion with the justifiable assumption that we have to deal with *elastic deformation* only. As early as 1941⁷ and 1942⁸ analyses have been made of the playback loss due to elastic deformation of

the modulated groove. A very elaborate investigation was conducted by Miller in 1950⁹; a few years later, Kantrowitz worked through Miller's analysis again, with special regard to vertical recording.¹² A different ap-



a.



b.

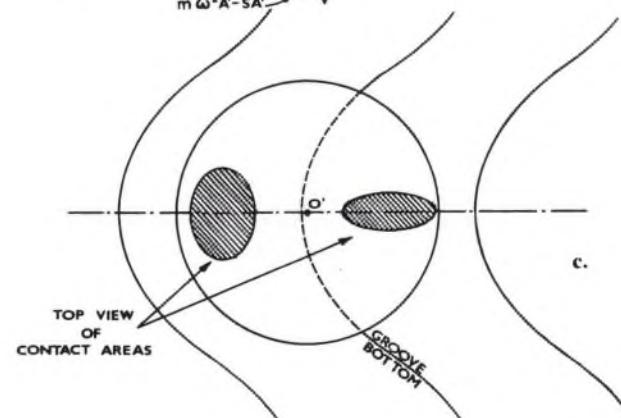


Fig. 3. a. Perspective view of a stylus tip (shown as a semi-sphere) sitting in the crest of a sinusoidally modulated groove. b. Vertical cross-section of a spherical stylus tip with the radius R , penetrating the walls of a modulated groove. The penetration into the convex wall is deeper than in the concave wall. c. Top view of the situation depicted in Figure 3b.

proach was given by Kerstens¹⁰ in 1956. The interested reader is referred to the relevant publications for more detailed information. In the following analysis a very much simplified approach is given, neglecting higher order components and tracing distortion.

When the tip of a stylus sitting on the crest of a sinusoidally modulated groove (see Fig. 3a) is moved in the lateral plane by the modulated groove, two lateral forces come into play; these are (see Fig. 3b):

1. A stiffness force SA' , where S is the suspension stiffness (or the inverse of the compliance) of the needle-armature and A' is the actual excursion, relative to the neutral line (silent groove), made by the tip.

2. An acceleration force $m\omega^2 A'$ required to move the dynamic mass m of the stylus-armature (assumed to be concentrated in the center of the stylus tip).

These forces oppose each other, so that the vertical stylus force F_v is altered by a lateral component ($m\omega^2 A' - SA'$). Decomposed components F_1 and F_2 exert forces, respectively, on the convex and the concave groove wall. Notwithstanding the fact that $F_2 > F_1$, the indentation in the concave wall will be less than that in the convex wall; this is because much more supporting area is presented by the concave wall to the sphere, so that a balance between the indenting force and the opposing force exerted by the groove wall is obtained with less penetration. Figure 3a makes it apparent that the concave wall "wraps around" the sphere, whereas the convex wall presents a "ridge" to it.

The net result is that the center of the sphere will be deflected from the ideal position (obtainable only with an infinitely stiff record) by an amount

$$\Delta = A - A' = (\delta_1 - \delta_2) \cos\beta. \quad (7)$$

The wall forces F_1 and F_2 can be expressed as follows:

$$F_{1,2} = \frac{F_v}{2\sin\beta} \pm \frac{(S-m\omega^2)A'}{2\cos\beta}. \quad (8)$$

Using the definition of K given in Eq. (4), the elastic deformation formula proposed by Hertz¹ becomes, in this case,

$$\delta = K \left[\frac{F_n^2}{R} \left(1 + \frac{R}{2\rho} \right) \right]^{\frac{1}{3}}. \quad (9)$$

The radius of curvature as calculated in Appendix II is:

$$\rho = V^2 / (\omega^2 A); \quad (10)$$

this, however, is the curvature in the lateral plane. In a plane perpendicular to a groove wall the radius of curvature equals:

$$\rho_{1,2} = \pm [V^2 / (\omega^2 A \cos\beta)] \quad (11)$$

In practice, the included groove angle is very close to 90° ; assume it to be exactly 90° in order to simplify the analysis to some extent. Substituting in Eq. (9),

$$(F_{1,2})^2 = \frac{F_v^2}{2} \left[1 \pm \frac{(S-m\omega^2)A'}{F_v} \right]^2 \text{ and } 2\rho_{1,2} = \pm \frac{2\sqrt{2}V^2}{\omega^2 A}$$

one obtains

$$\delta_{1,2} = K \left(\frac{F_v^2}{2R} \right)^{\frac{1}{3}} \left(1 \pm \frac{(S-m\omega^2)A'}{F_v} \right)^{\frac{2}{3}} \left(1 \pm \frac{R\omega^2 A}{2\sqrt{2}V^2} \right)^{\frac{1}{3}}. \quad (12)$$

At this point replace S by:

$$S = m\omega_0^2, \quad (13)$$

where ω_0 is the free resonance of the pickup system. Equation (7) can then be rewritten as follows:

$$\Delta = \frac{K}{\sqrt{2}} \left(\frac{F_v^2}{2R} \right)^{\frac{1}{3}} \left\{ \left[1 + \frac{m(\omega_0^2 - \omega^2)A'}{F_v} \right]^{\frac{2}{3}} \left(1 + \frac{R\omega^2 A}{2\sqrt{2}V^2} \right)^{\frac{1}{3}} - \left[1 - \frac{m(\omega_0^2 - \omega^2)A'}{F_v} \right]^{\frac{2}{3}} \left(1 - \frac{R\omega^2 A}{2\sqrt{2}V^2} \right)^{\frac{1}{3}} \right\}. \quad (14)$$

This expression can be simplified, since only the high frequency region is of interest, so that $\omega_0 < \omega$ (f_0 is usually around 1500 Hz for most pickups); thus, $\omega_0^2 \ll \omega^2$. Further applying the binomial theorem (introducing an error of less than 5% in the worst case):

$$\frac{R\omega^2 A}{2\sqrt{2}V^2} < 1, \text{ therefore} \\ \left(1 \pm \frac{R\omega^2 A}{2\sqrt{2}V^2} \right)^{\frac{1}{3}} \approx 1 \pm \frac{R\omega^2 A}{6\sqrt{2}V^2}$$

and

$$\frac{m(\omega_0^2 - \omega^2)A'}{F_v} \ll 1, \text{ therefore} \\ \left(1 \mp \frac{m\omega^2 A'}{F_v} \right)^{\frac{2}{3}} \approx 1 \mp \frac{2m\omega^2 A'}{3F_v}.$$

The lateral displacement of the sphere's center can now be written as:

$$\Delta \approx \frac{K}{\sqrt{2}} \left(\frac{F_v^2}{2R} \right)^{\frac{1}{3}} \left[\frac{R\omega^2 A}{3\sqrt{2}V^2} - \frac{4m\omega^2 A'}{3F_v} \right]. \quad (15)$$

The actual amplitude of excursion is thus, after some rearranging, found to be:

$$A' = A - \Delta \quad (16)$$

$$= A \cdot \frac{[1 - K(F_v^2/2R)^{\frac{1}{3}}(R\omega^2/6V^2)]}{[1 - K(F_v^2/2R)^{\frac{1}{3}}(2\sqrt{2}m\omega^2/3F_v)]}.$$

The derivation of this equation involves no mechanical damping, either in the pickup system proper or manifested as viscous losses in the record material. Consequently, we may find a situation where the actual amplitude A' is infinitely high, i.e., poles of A' (denominator of the fractional portion of Eq. (16) equals zero) and

a situation where the actual amplitude A' is zero, i.e., zeros of A' (numerator equals zero). The numerator is generally called the translation loss function; the denominator is termed the stylus/groove resonance function.

STYLUS/GROOVE RESONANCE

Equating the denominator in Eq. (16) to zero, and substituting K from Eq. (4), one obtains:

$$\omega_r^2 = \frac{3(2F_vR)^{\frac{1}{3}}}{2\sqrt{2}mK} = \frac{(96E_0^2F_vR)^{\frac{1}{3}}}{2\sqrt{2}m} \quad (17a)$$

or

$$f_r = [0.6362/\pi\sqrt{m}](E_0^2F_vR)^{\frac{1}{6}} \quad (17b)$$

where f_r = stylus/groove resonant frequency.

The stylus/groove resonance may also be expressed as follows:

$$f_r = 1/[2\pi\sqrt{mC}] = \sigma^{\frac{1}{3}}/2\pi m^{\frac{1}{2}}$$

where C = record compliance and σ = record stiffness, from which we obtain:

$$C = 0.617(E_0^2F_vR)^{-\frac{1}{3}} \quad (18a)$$

and

$$\sigma = 1.619(E_0^2F_vR)^{\frac{1}{3}}. \quad (18b)$$

From Eq. (17b) it is seen that the stylus/groove resonance frequency varies with the $\frac{1}{3}$ th power of stylus force F_v , all other parameters being held constant:

$$f_r \propto F_v^{\frac{1}{6}}.$$

Table I is computed using this expression; it also shows the magnitude of the record compliance C based on a tip radius $R = 17.8 \mu$ (0.7×10^{-3} in.) and a Vinylite record with $E_0 = 3.76 \times 10^9 \text{ N/m}^2$ (5.45×10^5 psi):

TRANSLATION LOSS

Equating the numerator in Eq. (16) to zero, we obtain:

$$\omega_c^2 = \frac{6V^2}{K} \left(\frac{2}{F_v^2 R^2} \right)^{\frac{1}{3}} = 12V^2 \left(\frac{2E_0}{3F_v R} \right)^{\frac{1}{3}} \quad (19a)$$

or

$$f_c = 1.513(V/\pi)(E_0/F_vR)^{\frac{1}{3}}. \quad (19b)$$

where f_c = the cut-off frequency at which the stylus does not move and hence no output from the pickup is obtained. This cutoff frequency varies inversely with the $\frac{1}{3}$ power of stylus force F_v :

$$f_c \propto F_v^{-\frac{1}{3}}.$$

Again, it is illustrative to compile the cutoff frequencies that will prevail when considering a 12-in. Vinylite record played with a given pickup at various bearing weights (Table II). The same tip radius as before is used ($R = 17.8 \times 10^{-6}$ m = 0.7×10^{-3} in.). In all cases the outer groove diameter (O.D.) equals 292 mm (11.5 in.) and the inner groove diameter (I.D.) equals

TABLE I. Record compliance in Terms of the stylus force.

Stylus Force F_v	Normalized f_r^*	Record Compliance
(Approx. grams)		(10^{-5} m/N)
10	1.3	2.16
8	1.255	2.33
6	1.2	2.56
4	1.12	2.96
2	1	3.68
1	0.89	4.6
0.5	0.795	5.92
0.1	0.607	10

100 mm (3.94 in.) for 78 and 45 rpm, 120 mm (4.7 in.) for $33\frac{1}{3}$ rpm, the standard diameters in the recording field.

RELATIVE PICKUP RESPONSE

Introducing ω_r from Eq. (17a) and ω_c from Eq. (19a) in Eq. (16), the relative pickup response may be written as

$$\frac{A'}{A} = \frac{1 - (\omega/\omega_c)^2}{1 - (\omega/\omega_r)^2} = \frac{1 + k_1 s^2}{1 + k_2 s^2} = G(s)H(s). \quad (20)$$

F. G. Miller⁹ evolved an expression for the function $H(\omega/\omega_r)$ which includes a damping factor

$$\epsilon = \zeta/m\omega_r$$

where ζ = damping coefficient of the combined system pickup and record (kg/sec):

$$H(\omega/\omega_r, \epsilon) = \frac{[1 - (\omega/\omega_r)^2 + \epsilon^2(\omega/\omega_r)^2]^{-\frac{1}{2}}}{\epsilon} \quad (21)$$

This equation is presented in graphical form in Fig. 4 with parameter $\epsilon = 0.1 \dots 2.0$. The damping factor ϵ tends to reduce f_r by a slight amount. Beyond the resonant frequency the H function approaches a decrease of 12 dB/octave asymptotically. This function may be recognized as a simple low-pass function having a pair of complex conjugate poles. Damping factor ϵ equals the inverse of the Q factor of the tuned circuit.

The function $G(\omega/\omega_c)$ is graphically depicted in Fig. 5a which will be discussed later on. Note that for $\omega > \omega_c$ the term reverses its sign; in other words, the stylus tip starts moving in a direction opposing the lateral groove excursions. Mathematically, A' may become larger than the original groove amplitude A . Not only would this be hard to visualize physically, but things become obscured because a third function (see section on scanning loss) comes into play, so that beyond the cutoff frequency A' will not become larger than A .

SCANNING LOSS

In the foregoing, the size of the contact surfaces between stylus tip and groove walls has been neglected.

* Based on the resonant frequency obtained with a given moving mass m at a bearing weight of 2 g. In other words, if a certain pickup exhibits a resonance of 50 kHz at a stylus force of 2 g, decreasing this force to 0.1 g (provided it can track at this low force) will bring down the resonance to $.607 \times 50 = 30$ kHz.

TABLE II. Cutoff frequencies with various bearing weights.

Turntable Speed (RPM)	Diameter	Groove Speed (m/sec)	Cutoff Frequencies in kHz at								Approximate Grams Newtons
			5 .05	4 .04	3 .03	2 .02	1 .01	.5 .005	.1 .001		
78	O.D.	1.20	93.5	101	111	127	160	202	345		
78	I.D.	.40	31	33.5	37	42.5	53.5	67	115		
45	O.D.	.69	54	58	64	73	92	116	198		
45	I.D.	.24	19	20	22	25.5	32	40	69		
33½	O.D.	.51	40	43	47	54	68	86	147		
33½	I.D.	.21	16.5	18	19.5	22	28	35	60		

Decreasing the stylus force (maintaining, of course, proper tracking abilities) moves the cutoff frequency *up* in the frequency spectrum and brings the resonant frequency *down* (Table I); it affects the cutoff frequency more than the resonant frequency.

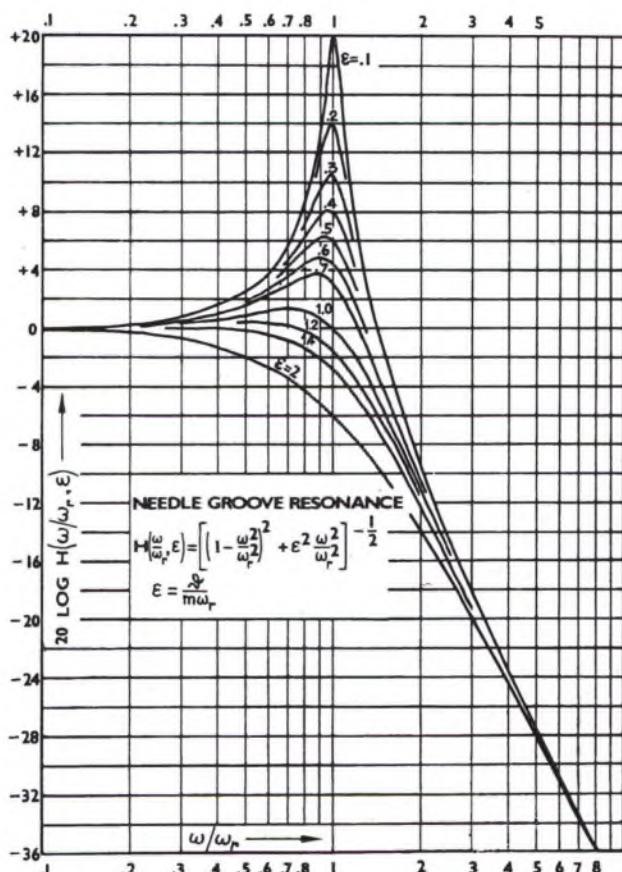


Fig. 4. Graphic presentation of the stylus/groove resonance function $H(\omega/\omega_r, \epsilon)$ for various values of damping parameter ϵ .

Should the recorded wavelength become comparable to the diameter of the contact surface, it is obvious that hardly any information can be "read" by the stylus tip. This phenomenon is termed the scanning loss and compares with the aperture loss encountered in optical and magnetic tape read-out systems. In the phonograph system, it is, however, of a more complex nature and it may partly offset the translation loss discussed in the previous sections. The two phenomena may probably not be simply added, because the scanning loss will tend to impart less amplitude information to the stylus tip and the groove wall forces F_1 and F_2 will differ less.

Miller's expression for the scanning loss function:

$$\begin{aligned} S\left(\frac{\pi d}{2\lambda}\right) &= S\left(\frac{\omega d}{4V}\right) = S(p) = \\ &= (2/\pi) \int (1-x^2)^{1/2} \cos 2px dx = \\ &= \sum_0^{\infty} \frac{p^{2n}}{(n!)^2} \cdot \frac{(-1)^n}{n+1}. \end{aligned} \quad (22)$$

The power series of this equation has its first zero for $\pi d/2\lambda = 1.9$, from which one obtains an indentation width

$$d = 3.8V/\pi f_s, \quad (23)$$

with f_s = scanning loss null frequency.

Introducing d as found in Eqs. (6) and (5), we write

$$3.8V/\pi f_s = 2K^{1/2}2^{-1/2}(F_v R)^{1/2}$$

from which

$$\omega_s = 3.8V \cdot 2^{1/2}/K^{1/2}(F_v R)^{1/2}. \quad (24)$$

Taking the ratio of scanning loss null frequency to the cutoff frequency defined in Eq. (19b), we find that this ratio is a constant = 1.55:

$$\omega_s/\omega_c = 3.8/\sqrt{6} = 1.55.$$

Both the function $G(\omega/\omega_c)$ and the function $S(\omega/\omega_s)$ are graphically presented in Figure 5a, the latter curve transposed for an ordinate scale $\omega/\omega_c = 1.55 \omega/\omega_s$.

For the fundamental of a recorded frequency, the S function may be neglected with little error, provided the recorded wavelength is large compared to the stylus/groove contact surface. As will be shown below, measured pickup response curves coincide with theoretical curves to a remarkably high degree. As has been pointed out earlier, these two functions probably may not be simply added. To illustrate the type of response that might result, however, the two effects are directly combined in one curve which is shown in Fig. 5b.

EXPERIMENTAL RESULTS

In order to check the validity of the theory under discussion, several experiments were carried out in which

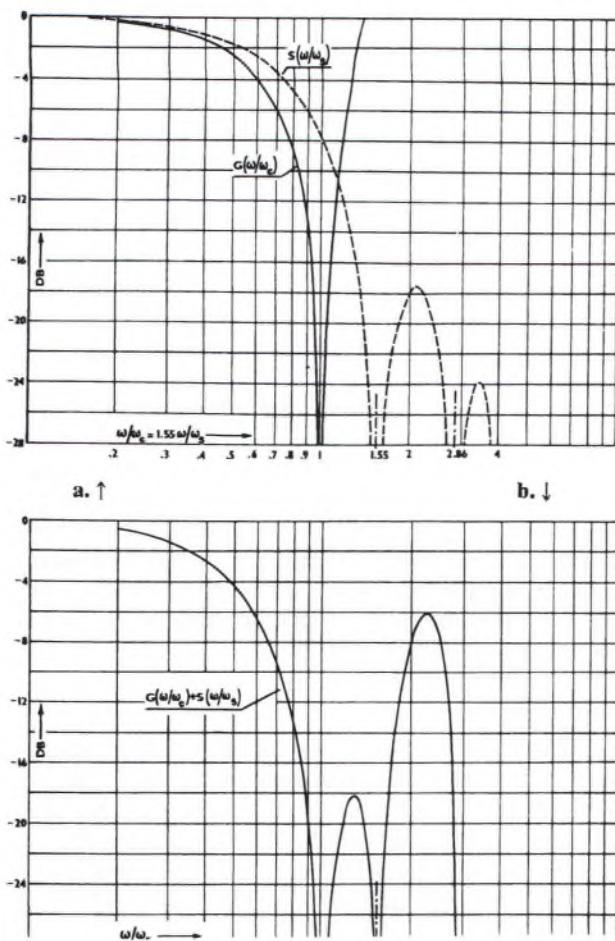


Fig. 5. a. Graphic presentation of cutoff function $G(\omega/\omega_c)$ and scanning loss function $S(\omega/\omega_c)$. b. Graphic presentation of combined cutoff (G) and scanning loss (S) functions.

measured frequency response curves were obtained and compared with the curves predictable through application of the theory.

Since commercial pickup cartridges are mostly heavily damped and their manufacturers are too often ignorant of the true value of moving mass, we used pickups of our own construction.

The performance of four different models with respect to the H and G function theories was examined. A value of $3.76 \times 10^9 \text{ N/m}^2$ for E_0 is used in all the following examples; all pickups were provided with a tip radius of 17.8μ (0.7×10^{-3} in.), and a stylus force of 1.5 g was used.

Pickup 1-3

This model had a moving mass (referred to the stylus tip) of 2 mg, and a damping factor $\epsilon = 0.3$.

Special frequency sweep records were cut with a constant stylus velocity of appropriate value, sweeping from 500 to 100,000 Hz. Subsequently, 12-in. Vinylite pressings were made; Young's modulus of the record compound used was not actually measured, but assumed to be $3.3 \times 10^9 \text{ N/m}^2$ (4.8×10^5 psi).

From Eq. (17b) it is found that $f_r = 17,800 \text{ Hz}$ and from Eq. (19b) $f_c = 58,000 \text{ Hz}$ for $V = 0.495 \text{ m/sec}$ ($33\frac{1}{3}$ rpm at approximately the outer diameter).

Figure 6a shows the H function plotted for $\epsilon = 0.3$, and also the G function. Furthermore, the two curves

are added together to obtain the dashed curve, which is then transferred to Fig. 6b where it is compared with the

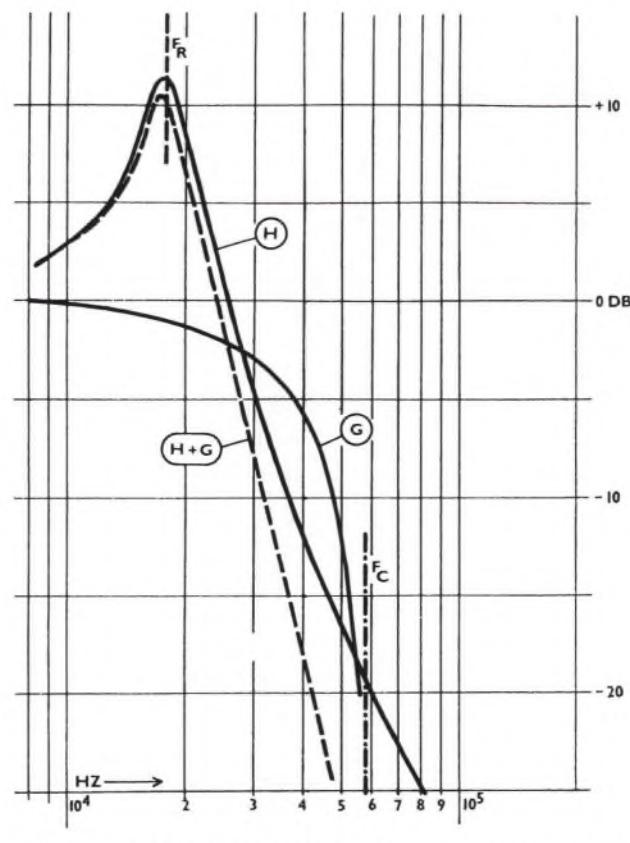


Fig. 6. a. Theoretical high-frequency response curve for model 1-3 pickup with moving mass $m=2 \text{ mg}$, damping $\epsilon=0.3$, tip radius $R=17.8 \mu$ and stylus force $F_r=1.5 \text{ g}$ on Vinylite with groove velocity $V=0.495 \text{ m/sec}$.

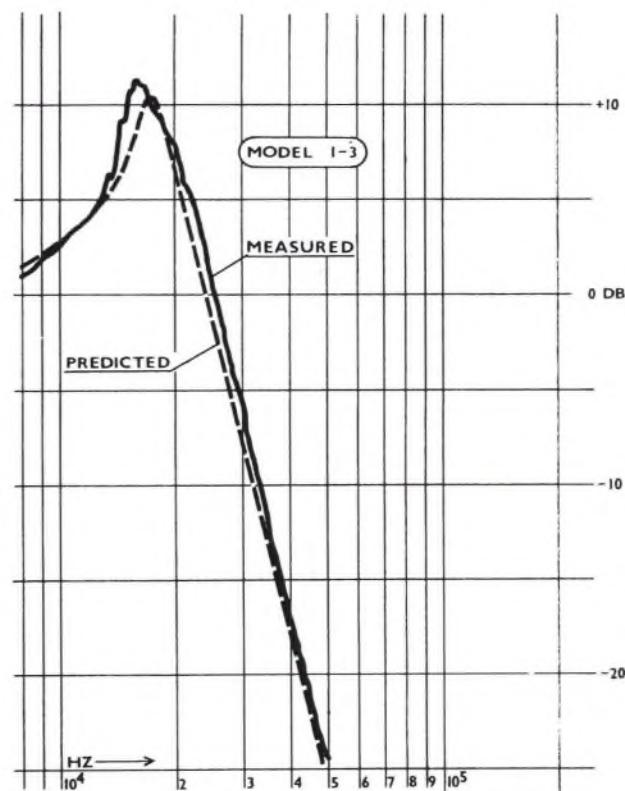


Fig. 6. b. Comparison of measured and theoretical response curves for model 1-3 pickup. Parameters as listed for Fig. 6a.

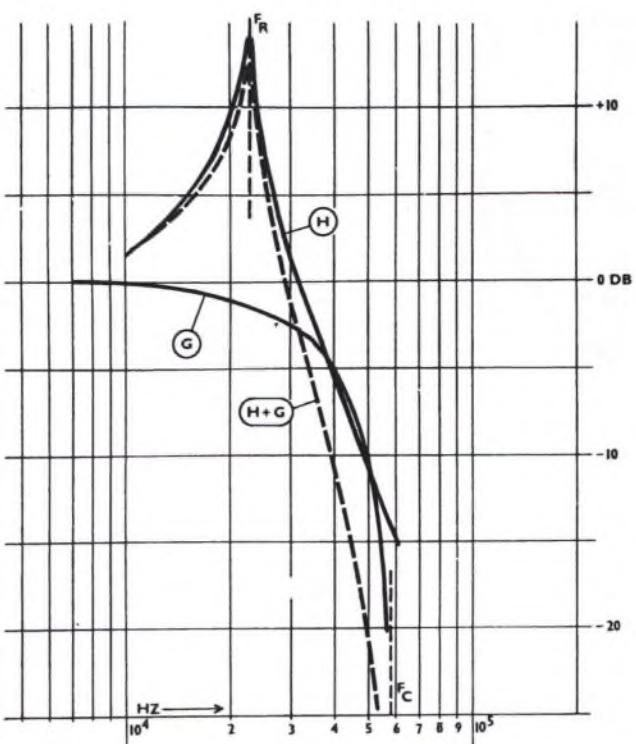


Fig. 7. a. Theoretical high frequency response curve for model 2-2 pickup with moving mass $m = 1.2$ mg, damping $\epsilon = 0.2$. Other parameters as in Figure 6a.

measured response of this particular pickup. Note that the two curves match within 2 dB or better.

Pickup 2-2

For this model, the moving mass was 1.2 mg and the damping factor $\epsilon = 0.2$.

Again, we calculate $f_r = 23,000$ Hz; since the stylus

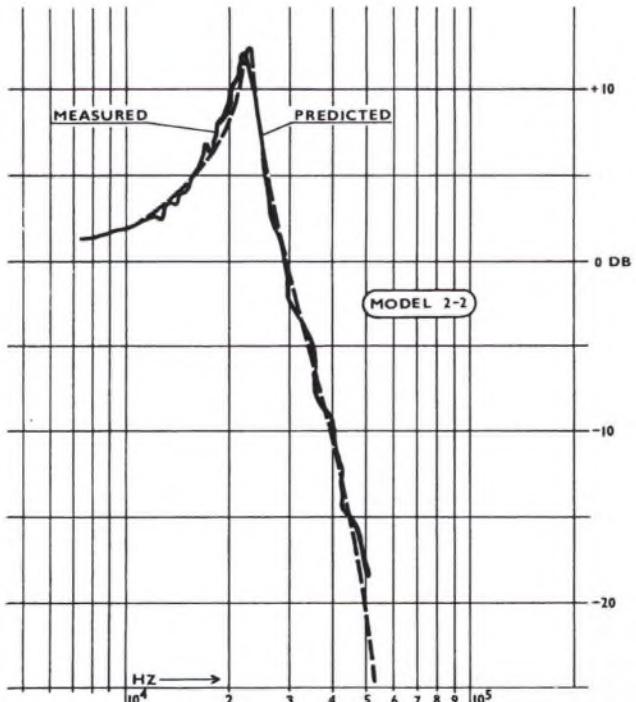


Fig. 7. b. Comparison of measured and theoretical response curves for model 2-2 pickup. Parameters as listed for Fig. 7a.

force, tip radius, groove speed and record material were the same as in the preceding example, we again have $f_c = 58,000$ Hz.

Both functions are plotted and added in Fig. 7a; the resonance peak is somewhat higher with this pickup because of a lower damping value. Again the $(H+G)$ plot is transferred to Fig. 7b and compared with the measured response, which it matches very closely.

Pickup 3E

This pickup has a very much lower moving mass and also more damping. Its moving mass = 0.4 mg and $\epsilon = 0.7$.

With the given characteristics, this pickup should exhibit a stylus/groove resonance $f_r = 40,000$ Hz. It was decided to assure a cutoff frequency well beyond this, so a higher groove speed (.79 m/sec) was used for playback. As a result, f_c should occur at 92.5 kHz. Both H and G functions are depicted in Fig. 8a and added to obtain the dashed curve. Comparison of the latter with the measured curve again discloses a very close match, within 2 dB (see Fig. 8b).

Pickup 5B-T

To investigate the pickup response under circumstances where $f_r > f_c$, this pickup was designed to have extremely low moving mass, while the stylus force was increased to 2 grams. Its moving mass was 140 μ g, and $\epsilon = 0.2$.

Figure 9a shows the H function for a calculated $f_r = 70,000$ Hz and the G function for $f_c = 122.5$ kHz at an

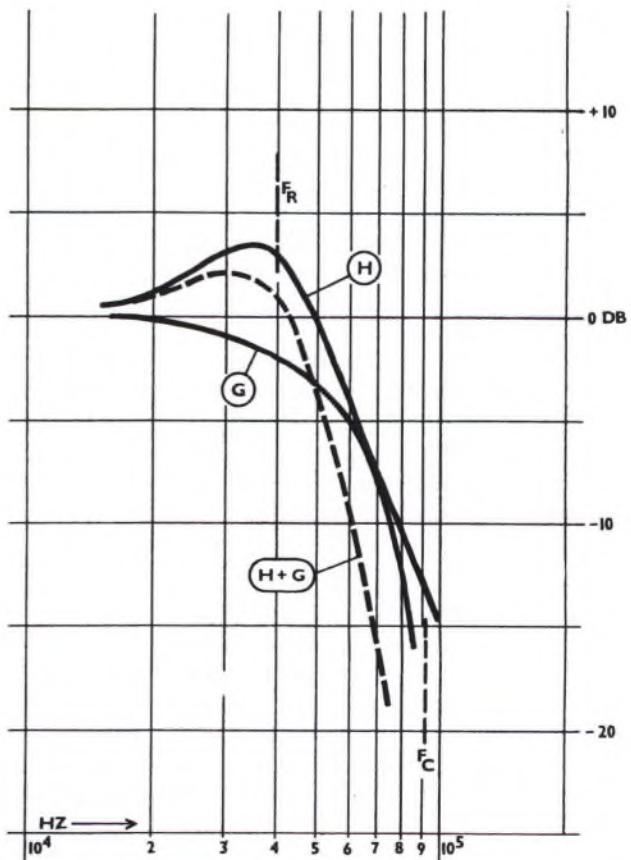


Fig. 8. a. Theoretical high frequency response curve for model 3E pickup with moving mass $m = 400$ μ g, damping $\epsilon = 0.7$. Other parameters as in Figure 6a, except for groove velocity $V = 0.79$ m/sec.

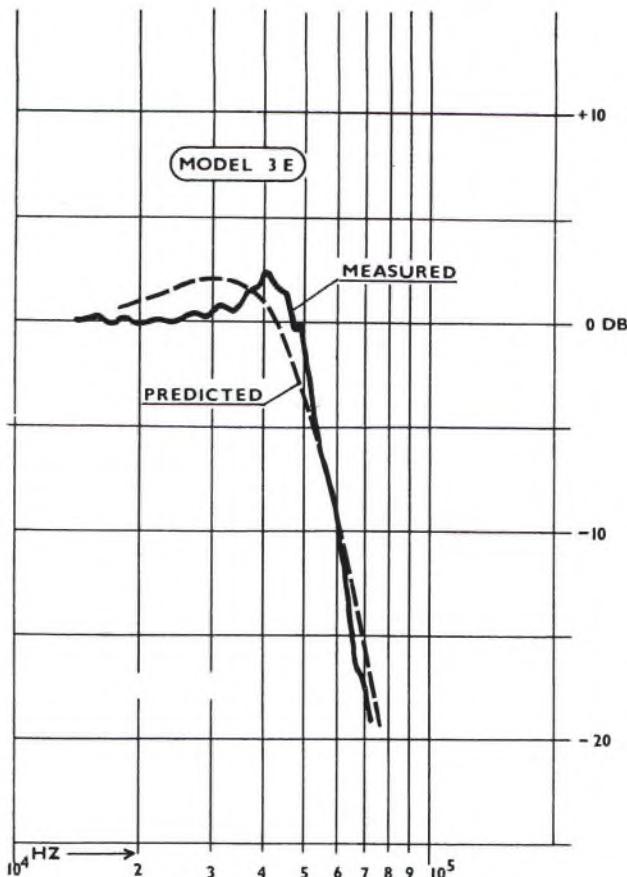


Fig. 8. b. Comparison of measured and theoretical response curves for model 3E pickup. Parameters as listed for Fig. 8a.

outer-diameter groove velocity of 1.15 m/sec. The resulting addition (dashed curve) is next transferred to Fig. 9b, where it agrees with the measured response within 2 dB (Curves I and II). Furthermore, the graph depicts the measured response (Curve III) at an inner-diameter groove velocity of 0.6 m/sec ($f_c = 64$ kHz). The difference between this inner diameter response and the H function (see Fig. 9a), which should normally "match" the G function for $f_c = 64$ kHz, has also been depicted. To avoid a cluttered graph this Curve IV has been transferred to Fig. 9c. The theoretical G function is shown as Curve V and it can be clearly seen that it does not match Curve IV. However, if the scanning loss function is added to the theoretical G curve, the response indicated by Curve VI is obtained, which is more of a resemblance to Curve IV. The conclusion is that the S function should not be ignored at very high frequencies. This is quite understandable since the recorded wavelengths in this case are indeed comparable to the finite size of the needle/groove contact surfaces! The fact that a very close match between theoretical and measured curves cannot be shown here is probably due to the possibility that the G and S functions cannot be simply added since these two phenomena are really two aspects of the same physical factors. Moreover, the radius of groove curvature at these frequencies does not satisfy the condition $\rho \geq 1.5R$.

CONCLUSION

The H , G and S functions prove to be valuable tools in predicting the performance of a pickup design. The

vertical stylus force must be kept low enough so as to cause mainly elastic deformation of the groove, at the same time assuring proper tracking. The G and S functions are the prime factors that limit the passband of a pickup/record playback system. An obvious improvement would be to use a record material that is very much stiffer than the Vinylite compound presently used. Preliminary tests with nickel records have not been successful, possibly due to the fact that almost complete plastic flow prevailed, even at a stylus force of 1 gram. Since the metal is "very much harder" ($E = 21 \times 10^{10}$ N/m² $\approx 30 \times 10^6$ psi), the indentation is indeed very small. The obvious result is that the pressure exerted by the stylus on the groove walls is very high and exceeds the yield point. The sharp knee in the stress/strain curve of nickel means that plastic flow sets in instantaneously, with hardly any elastic components.

These experiments indicate that, at the present state of the art, only a very much lower stylus force offers a solution for increasing the passband of a pickup. This, however, can be possible only if the moving mass is small enough to cope with the extremely high accelerations encountered in the high-frequency region. However, a record material that is very stiff and features a very high yield point, so that deformation of the groove walls is limited to the elastic region with very little indentation, would indeed be a worthwhile move towards a truly wideband phonograph pickup system.

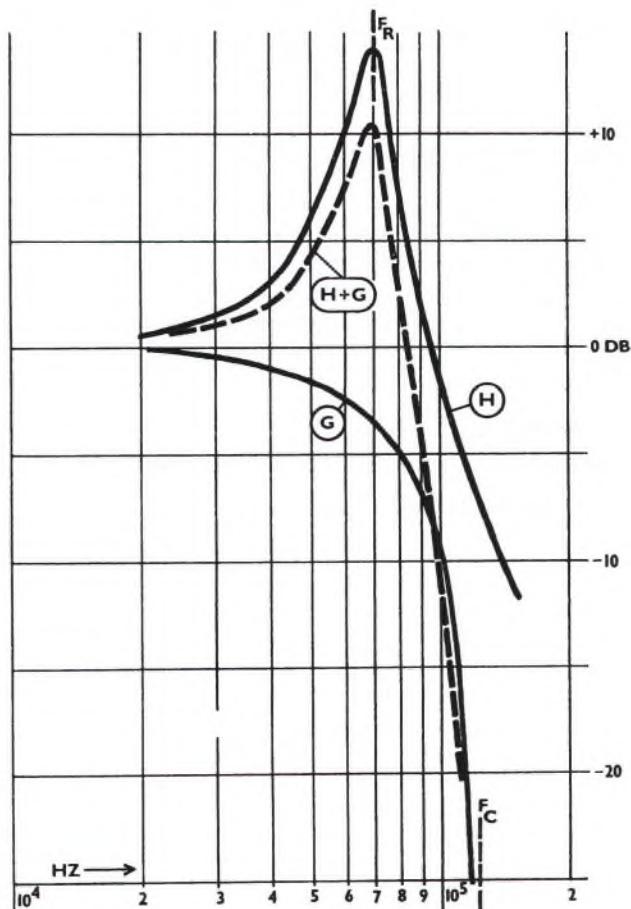


Fig. 9. a. Theoretical high frequency response curve for model 5B-T pickup with moving mass $m = 140 \mu\text{g}$, damping $\epsilon = 0.2$, tip radius $R = 17.8 \mu$ and stylus force $F_v = 2 \text{ g}$ on Vinylite with a groove velocity $V = 1.15 \text{ m/sec}$.

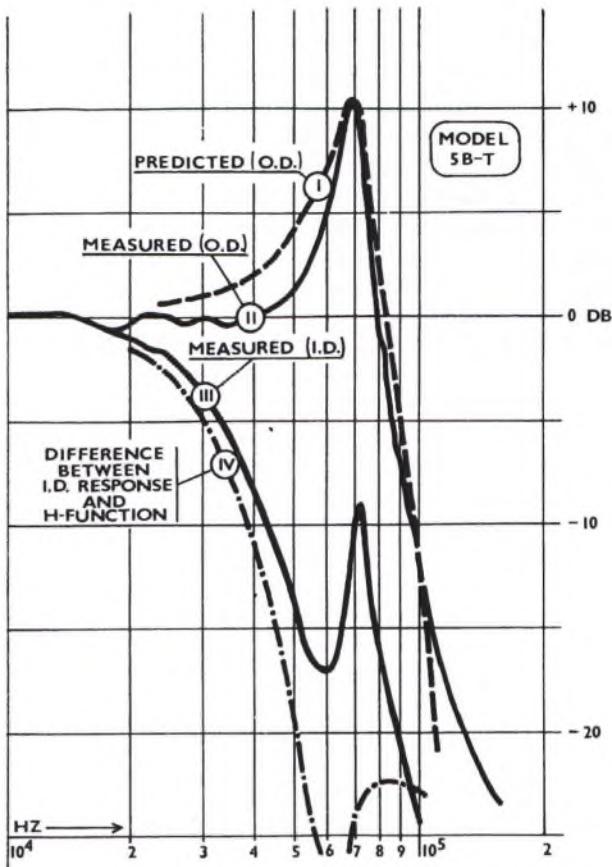


Fig. 9. b. Comparison of predicted (Curve I) and measured (Curve II) response curves for model 5B-T pickup. Parameters as in Figure 9a. Curve III is the measured response curve for $V=0.6$ m/sec, and Curve IV depicts the difference between H function and Curve III.

ACKNOWLEDGEMENTS

The author wishes to acknowledge the cooperation of Romolo and Richard Marcucci and Sal Gualtieri of Capps and Company, Inc. who supplied the special stylus tips for the experimental pickups and by H. Bregman of Sonic Recording Products, Inc. who did such an excellent job in the processing of the many test records.

APPENDIX I Summary of Symbols and Terms Used

(In Alphabetical Order)

Symbol	Term	Dimension (Rationalized MKS)
A	Recorded amplitude	m, meter
A'	Stylus excursion	m
C	Compliance of record material = $1/\sigma$	m/N , meter per Newton
d	Indentation width	m
E	Young's modulus of elasticity	N/m^2 , Newton per square meter
E_0	Constrained Young's modulus = $E/(1-\nu^2)$	N/m^2
F	Force	N, Newton
F_n	Force, normal to a surface	N
F_v	Vertical stylus force	N
f	Recorded frequency	Hz, Hertz
K	Materials constant	$(m^2/N)^{2/3}$

m	Dynamic moving mass of armature	kg, kilogram
R	Radius of stylus tip	m
S	Stiffness of armature suspension	N/m
V	Tangential groove velocity	m/sec, meters per second
ω	Recorded frequency = $2\pi f$	rad/sec, radians per second
ω_c	Cutoff frequency	rad/sec
ω_0	Free armature resonance of pick-up	rad/sec
ω_r	Stylus/groove resonance frequency	rad/sec
ω_s	First scanning loss null frequency	rad/sec
β	Half included groove bottom angle	degrees
δ	Indentation depth	m
δ_0	Indentation depth in a silent groove	m
Δ	Difference between A and A'	m
ϵ	Damping factor = $\zeta/(m\omega_r)$	(dimensionless)
ζ	Damping coefficient	kg/sec, kilogram per second
λ	Recorded wavelength	m
ν	Poisson's ratio	(dimensionless)
ρ	Radius of curvature of recorded waveform	m
σ	Stiffness of record material	N/m

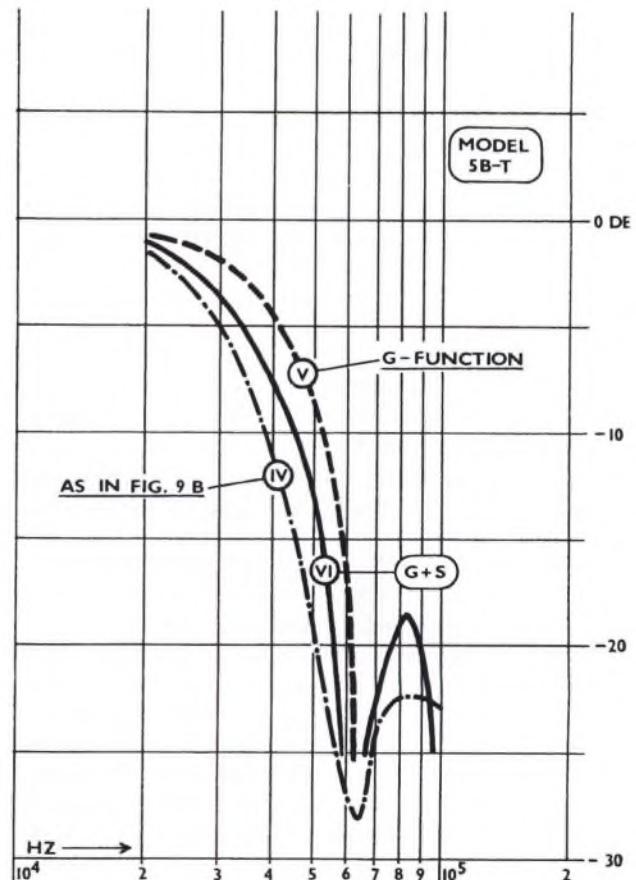


Fig. 9. c. Situation at very high frequencies where the scanning loss function cannot be ignored. Curve IV is the same as Curve IV in Figure 9b and Curve V is the theoretical G function, whereas Curve VI is the combined G and S function. Model 5B-T pickup, parameters as listed for Figure 9a, except groove velocity $V=0.6$ m/sec.

APPENDIX II

Calculation of the Radius of Curvature

The radius of curvature at any point of a function $f(y)$ is given¹¹ by:

$$\rho = (1+y'^2)^{3/2}/y'' \quad (\text{A1})$$

where y' stands for the first derivative and y'' is the second derivative of the function $y = f(x)$.

In the case of a sinusoidal groove modulation (see Fig. A1)

$$y = A \sin \omega t = A \sin(2\pi x/\lambda) \quad (\text{A2})$$

where λ is the wavelength equal to the tangential groove velocity V divided by the frequency f :

$$\lambda = V/f. \quad (\text{A3})$$

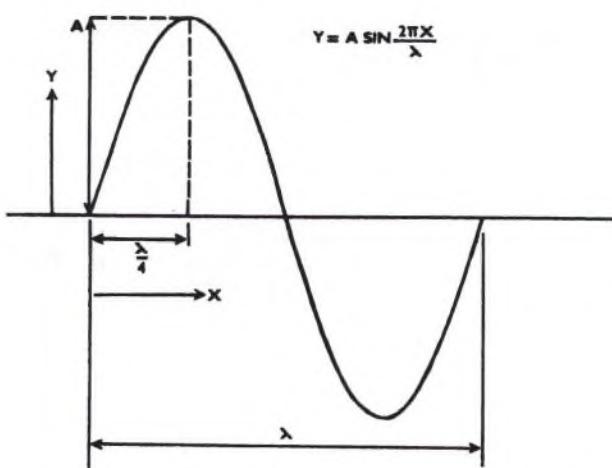


Fig. A1. Sinusoidal groove modulation with amplitude A and wavelength λ .

The first derivative $y' = (2\pi A/\lambda) \cos(2\pi x/\lambda)$, and the second derivative $y'' = -(4\pi^2 A/\lambda^2) \sin(2\pi x/\lambda)$. Equation (A1) then is written as:

$$\rho = \frac{[1 + (4\pi^2 A^2/\lambda^2) \cos^2(2\pi x/\lambda)]^{3/2}}{(4\pi^2 A/\lambda^2) \sin(2\pi x/\lambda)}. \quad (\text{A4})$$

It is obvious that a minimum for ρ occurs at $x = \lambda/4$, $3\lambda/4$, etc. Substituting $x = \lambda/4$ in Eq. (A4) gives

$$\rho = \lambda^2/4\pi^2 A. \quad (\text{A5})$$

Replacing λ by V/f , Eq. (A5) becomes Eq. (10) in the text:

$$\rho = V^2/4\pi^2 f^2 A = V^2/\omega^2 A. \quad (\text{10})$$

REFERENCES

1. H. Hertz, *Gesammelte Werke*, vol. 3, p. 155 (1894).
2. F. V. Hunt, "Elastic-Plastic Instability Caused by the Size Effect and Its Influence on Rubbing Wear," *J. Appl. Phys.* **26**, 850 (1955). See also "On Stylus Wear and Surface Noise in Phonograph Playback Systems," *J. Audio Eng. Soc.* **3**, 2 (1955).
3. D. A. Barlow, "Comments on the Paper, 'On Stylus Wear and Surface Noise,'" *J. Audio Eng. Soc.* **4**, 116 (1956).
4. J. Walton, "Gramophone Record Deformation," *Wireless World* **67**, 353 (1961).
5. R. M. Davies, "The Determination of Static and Dynamic Field Stresses Using a Steel Ball," *Proc. Roy. Soc. London A* **197**, 416 (1949).
6. R. A. Walkling, "Dynamic Measurement of the Hardness of Plastics," Doctoral Dissertation, Acoustics Research Laboratory, Harvard University, Cambridge, Mass. (May 1963).
7. O. Kornei, "On the Playback Loss in the Reproduction of Phonograph Records," *J. Soc. Motion Picture Engrs.* **37**, 569 (1941).
8. S. J. Begun and T. E. Lynch, "The Correlation Between Elastic Deformation and Vertical Forces in Lateral Recording," *J. Acoust. Soc. Am.* **13**, 284 (1942).
9. F. G. Miller, "Stylus Groove Relations in Phonograph Records," Doctoral Dissertation, Harvard University, Cambridge, Mass. (March 1950).
10. J. B. S. M. Kerstens, "Mechanical Phenomena in Gramophone Pickups at High Audio Frequencies," *Philips Technical Review* **18**, 89 (1956/57).
11. G. E. F. Sherwood and A. E. Taylor, *Calculus*, (Prentice-Hall, Englewood Cliffs, N. J., 1954).
12. P. Kantrowitz, "High-Frequency Stylus-Groove Relationships in Phonograph Cartridge Transducers," *J. Audio Eng. Soc.* **11** (1963).
13. T. Shiga, "Deformation Distortion in Disc Records," *J. Audio Eng. Soc.* **14**, 208 (1966).

THE AUTHOR



C. R. Bastiaans was born in the former Netherlands East Indies in 1924. He attended school in Holland where he received the diplomas of radio engineer and electrical engineer. In 1950, after five years as an officer in the army, he joined the Philips telecommunication department as a technical-commercial engineer. In 1957, he became head of the electro-mechanical development group of Philips Phonographic Industries. Following this, in 1963, he joined the research staff at

Westinghouse in Pittsburgh, Pennsylvania where he is in charge of Special Projects in their Research and Development Center.

Applications have been made for several of Mr. Bastiaans' patents in the audio field. He has been a contributor to the Philips Technical Review and has been published previously in the Audio Engineering Society Journal. Mr. Bastiaans is a member of the Audio Engineering Society.

COMMENTS ON THE STYLUS GROOVE RELATIONSHIP IN PHONOGRAPH PLAYBACK SYSTEMS

D. A. BARLOW

Uxbridge, Middlesex, England

THE paper by C. R. Bastiaans entitled "Factors Affecting the Stylus/Groove Relationship in Phonograph Playback Systems" (*J.A.E.S.* **15**, 389, Oct. 1967) may be supplemented by the following comments.

Some years ago, Hunt¹ noted from the Hertz equations that for a pick-up stylus on vinyl, the yield condition is reached at loads of only a few milligrams. Beyond this, these equations cannot be used, as by definition they apply only to the fully elastic range. The test results given by Mr. Bastiaans in his Fig. 2 are not below

any "theoretical limit", as there is as yet no theory that applies beyond the fully elastic range. The writer has described elsewhere² the behavior of a material under an indenter. The slope of the log impression diameter *vs* log load curve is $\frac{1}{3}$ in the fully elastic region, increasing to about $\frac{1}{2}$ in the fully plastic range. The scatter of the results is such that lines of this slope could equally well be drawn through the points. Thus there is nothing unusual about the slope of the test results and this is not connected with the elastic springback of Hunt's indentations.

Nevertheless, the points should lie above the line extrapolated from the elastic range, and the two styli should give similar track widths for a given load in this range. An incorrect value may have been taken for the modulus and/or the material may show a skin effect. Mr. Bastiaans states that the results varied around the record. Walkling³ also found this and showed that it was due to varying residual

stresses in the record, due doubtless to rapid irregular cooling from molding. Rapid cooling gives residual compression in the surface layers and residual tension in the interior of the material. The effect of residual compression would be to increase the apparent hardness, which would vary with the depth of the impression. If consistent results are to be obtained, slow-cooled stress-free material must be used. Where a suitable controlled stress system can be produced in a material, improved properties may result; toughened glass is an example of this. A consistent residual stress system appears to be difficult to obtain in modern records, judging by the amount of warping.

REFERENCES

1. F. V. Hunt, *J. Audio Eng. Soc.* **3**, 1 (1955).
2. D. A. Barlow, *Wireless World* **70**, 4 (1964).
3. R. A. Walkling, *O.N.R. Tech. Memo No. 49* (1963).

On the Interaction between Tracing Correction and a Bandwidth Limitation

DUANE H. COOPER

University of Illinois, Urbana, Illinois

A double-bandwidth requirement proposed by Hunt as necessary to the making and reproducing of tracing-corrected recordings is shown to be irrelevant to the reproducing element by an argument using a factoring demonstrated by Miller. The implications of Miller factoring in explaining the residual distortion in tracing-corrected recording as being essentially the sole result of an ameliorated deformation error are sketched. Analysis of numerical examples of interactions between tracing correction and bandwidth limitations for the recording element shows the interaction to be of negligible consequence for practical cases. An argument is made that the full realization of the growth potential of disc-recording technology requires tracing error to be eliminated through a correct shaping of the groovewall waveform to favor the use of low-deformation styli in playback.

INTRODUCTION Advocating the use of tracing-error correction by means of disc-to-disc dubbing with inversion of polarity [1], F. V. Hunt nevertheless asserted [2] that "the passband of both the rerecording and the final reproducing element must extend to an upper limit of $2f_u$ in order to achieve the desired cancellation of distortion at all frequencies below f_u ." Though the intention was obviously to justify the laudable goal of obtaining as great a bandwidth as possible for such elements, it is possible that the assertion goes too far. If widely taken literally, it would have had the effect of blunting an advocacy for both the dubbing method as well as for electronic shaping [3-6] of the waveform to be recorded. The requirements applicable to one would be applicable to the other, and it could be thought that the use of tracing correction by either method ought to be deferred until the attainment of the double bandwidth were practical. Though an effective doubling of bandwidth in the recording process is possible through half-speed recording, the tedium of that, though not as unattractive as disc-to-disc dubbing, combined with the difficulty of maintaining good low-frequency response, is not likely to

win many advocates. Also, the availability of cutters with a 30 kHz bandwidth, for example, may not be predicted for the near future.

It is the purpose of this paper to inquire into the validity of the double-bandwidth requirement in its applicability to both the recording and reproducing elements. Taken literally, the statement of this requirement seems to imply both a necessity and a sufficiency, namely, that the desired cancellation is not obtained unless the requirement is met, while, on the other hand, a meeting of this requirement will provide for as complete a cancellation as may be desired, insofar as any bandwidth limitation may be involved. The plausibility for the necessity seems to rest, at first thought, upon the recognition of tracing-error generation as involving a second-order nonlinearity and upon an intuitive equating of "second order" with "second harmonic." This, however, is too naive to stand examination, and it should not be supposed that such a justification was ever intended. What is needed is a detailed accounting of how the distortion components at frequencies below f_u , many of which are generated as intermodulation products from

original components in that same frequency range, depend for their cancellation upon the presence of components whose frequencies lie between f_u and $2f_u$, many of which also are intermodulation products. Though it is now possible to give such an accounting, it is a complicated matter to manage the details properly.

The sufficiency of the double-bandwidth requirement may also be called into question. The tracing-error non-linearity presents itself not only to second order, but also, though more weakly, to third and higher orders as well. Why should there not be a $3f_u$ requirement or one specified by $4f_u$, etc.? Generally, it will be true that the more nearly perfect cancellation would obtain for the wider bandwidths.

It may be fairly objected that the double-bandwidth limitation as stated should not be taken too literally, that it stopped short of specifying greater bandwidths, because, as a practical matter of engineering judgment, the major part of the cancellation would obtain for a mere doubling of bandwidth, and that further bandwidth extension would result in further improvement of only a marginal character. Possibly also, the cancellation was being compared to that obtaining for push-pull systems for which, unlike tracing correction, cancellation obtains only for even-order terms. Rather than make interpretive judgments on a conjectural basis, however, it is better to proceed with the purpose of this paper, namely the inquiry into the factual basis. For this purpose, a literal interpretation, though possibly unfair, will be helpful in sorting out the relevance of various parts of the statement so that attention may then be given to the more practical engineering considerations.

The applicability of the double-bandwidth requirement to the final reproducing element is examined by seeking to pinpoint the step in the reproducing sequence at which the cancellation would occur. With the help of a factoring first demonstrated by F. G. Miller [7], it is shown that the cancellation occurs prior to the action of any frequency-response-determining element in reproduction, so that no bandwidth considerations involved in the final reproducing element are relevant to achieving the desired cancellation. In the case of the dubbing method, the intermediate reproducing element could be lumped with the recording elements at the price of introducing complications in the analysis of dubbing systems that place their further study beyond the scope of this paper. The relevance of the recording bandwidth is then to be examined for a possible explanation of the residual distortion measured in experimental demonstrations [3] of tracing-error correction.

Since the demonstration of tracing-error correction obeyed the double-bandwidth requirement, its sufficiency is temporarily assumed, pending subsequent investigation, so that the wider implications of Miller factoring may be sketched. In view of this factoring, it is shown that the residual distortion, if not caused by a bandwidth limitation, must be essentially pure deformation-error distortion. It is demonstrated that, apart from causing the appearance of this distortion in pure form for the first time, the removal of tracing error has greatly reduced the severity of the deformation as well. Other aspects of the interaction between tracing error and deformation error are briefly discussed, including appar-

ently fortuitous observations [8] of a tendency for the two to cancel one another.

The applicability of the double-bandwidth requirement to the recording element is examined by means of numerical examples. These examples are amenable to theoretical calculation using formulas given in the Appendix. Distortion components in the frequency range below $2f_u$ are calculated as generated by the complement of tracing error from a pair of tones with frequencies below f_u . The two cases, containing distortion components below f_u and below $2f_u$, are then submitted to the tracing-error process and all relevant combination tones are computed. It is seen that cancellations do obtain for the below f_u distortion components, which cancellations are assisted by combination tones arising from the components in the f_u to $2f_u$ range. These latter combinations, however, are seen to be of third order in comparison to the first-order generation from the below f_u components. Thus, even for very moderate reductions in signal level (below curvature overload), the contributions to the cancellation from components lying above f_u are essentially negligible. The conclusion is that, while the interaction of tracing-error correction with a bandwidth limitation is real in principle, it is inconsequential in practice, and the double-bandwidth requirement may be dismissed from further consideration, to be replaced by a quite permissive level requirement.

The concluding remarks are addressed to the consequences of the showing that the residual distortion seen in demonstrations of tracing-error correction is uncontaminated by any effects due to an interaction with a bandwidth limitation, so that, because of Miller factoring, this residuum must be attributed to nothing else, apart from obviously identifiable artifacts, than deformation error. This observation, coupled with the observation that the removal of tracing error accomplishes, via the tracing-deformation interaction, the removal of a large part of the deformation-error also, provides a glimpse of the growth potential inherent in the disc-recording art. The future of the art appears to lie with making tracing error and deformation error independently accessible to control. It is a future in which the reduction of one error at the expense of augmenting the other by invoking, for example, an increase in sharpness of stylus curvature, will be seen to be an obsolescent practice by contrast with a correct shaping of the groovewall waveform favoring the use of low-deformation playback styli.

RELEVANCE TO THE FINAL REPRODUCING BANDWIDTH

The double-bandwidth requirement clearly should not be understood to be relevant to a bandwidth limitation imposed by any part of the reproducing system that receives its signal from the stylus-groovewall interaction. Thus, to take an extreme example, the notion of the "reproducing element" should not extend to the inclusion of a reproducing loudspeaker as part of that element. Any cancellation that would obtain by tracing correction would have been completed, to whatever extent possible, long before the signal exciting the loudspeaker was finally formed. Thus, the bandwidth limitation imposed by such an element is irrelevant to the adequacy of such

cancellation. Similarly, moving a step closer to the root of the problem, it may be seen that the frequency response of the electromechanical transducer driven by the stylus motion could also be irrelevant to the adequacy of the cancellation. That is to say, if the motion of the stylus tip were somehow free of any frequency limitation, then that motion must be uniquely determined by the stylus-groovewall interaction, and, regardless of how that motion is translated into a transducer action, the cancellation will already have been represented in that motion, so that the style of the transducer action would be irrelevant to determining its adequacy.

The style of analysis implicit in the preceding paragraph is one in which the overall transformation of the signal from the stylus-groovewall interaction to its ultimate point of emergence from the reproducing system is to be factored into a sequence of transformations. Then, since the complementary distortion components "planted" in the groovewall waveform are intended to cancel those arising directly from the stylus-groovewall interaction, the transformations invoked at later points in the sequence can have no bearing on the adequacy of that cancellation. The root of the problem lies in determining the exact point at which the factoring is still possible but the cancellation has been completed, as nearly as ever it will be, prior to that point. In the preceding paragraph, the illustrative example of "factoring off" the loudspeaker is obviously correct, while the similar factoring of the transducer action involves an "if" which is not so obviously justified. There could, it would seem, actually be a frequency limitation in the motion of the stylus that would be imposed by the exact manner of the transducer action. A more penetrating analysis is needed if this factoring is to be successful.

The more penetrating analysis is made possible through a radical shift of viewpoint taken by F. G. Miller [7], one of Hunt's students, in realizing that the groovewall deformation may be represented as deviations from an undeformed stylus trajectory. The superposition of these deviations upon the undeformed stylus trajectory, then, forms the true stylus trajectory, i.e., the deformed stylus trajectory. These deviations are properly identified with the actual groovewall deformation. The concept of the undeformed stylus trajectory is that of the trajectory that would be followed in the rigid-groovewall case, and was entirely specified in terms of the groovewall waveform by the tracing-error transformation. For this, Miller used the correct power-series representations of Lewis and Hunt [9]. By choosing a coordinate system moving along the undeformed trajectory, Miller was able to identify the deformation force as the difference between a force of constraint for that trajectory and the actual force on the stylus. In this way he obtained differential equations for the deformation, which he then solved by making perturbation expansions in trigonometric series for a sinusoidal groovewall waveform, assuming the Hertzian deformation law.

Miller's shift in viewpoint is radical in the true sense, because it strikes to the root of the factoring problem. His handling of the dynamics of the deformation process shows that it is perfectly valid to regard the undeformed trajectory as the input to the dynamic system which then responds by producing the deformed trajectory.

This achieves the factoring of the stylus-groovewall interaction into two parts. First there is the transformation of the groovewall waveform into the undeformed stylus trajectory by the tracing-error transformation as if the groovewalls were rigid; then there follows the generation of deformation error to produce the deformed trajectory. This factoring is shown in the block diagram of Fig. 1.

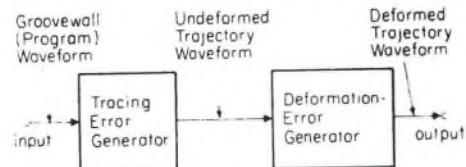


Fig. 1. Miller factoring of the stylus-groovewall interaction.

Though flawed in its execution, Miller's approach to the dynamics is undoubtedly correct. The flaw in execution consisted in taking the deformation to lie only in the direction of the normal to the quiescent groovewall rather than along the instantaneous normal to the undeformed wall. Thus, the delay-modulating character of the deformation was overlooked. Also, the groovewall curvature, a parameter in the Hertzian formula, was taken to be the value not at the contact but at a point lying below the stylus center. These flaws were not actually inherent in Miller's approach, since the correct direction of the deformation could have been specified in terms of the normal to the undeformed trajectory, and the curvature at the contact could also have been specified in terms of the curvature of the undeformed trajectory. The curvatures in the Hertzian theory are specified to be those of the undeformed bodies at the initial point of contact [10]. Thus, the undeformed trajectory could have been taken as providing the full input specification; this will always be the case, since a unique feature-for-feature geometrical correspondence exists between the groovewall waveform and the undeformed trajectory.

In a more recent paper, Shiga [11] took the correct direction for the groovewall deformation, so that his analysis, besides being more comprehensive, is the more definitive treatment within the Hertzian theory. He eventually used the same curvature approximation as Miller's, but he was able to justify the approximation. The Miller factoring is not necessary, and Shiga did not use it. Nevertheless, Shiga's analysis admits of Miller factoring. In Shiga's fundamental Eq. (20), for example, the terms may be identified as arising solely from tracing error, solely from deformation error, or from the two jointly. Shiga showed how the classic tracing-error transformation obtains if the deformation is set to zero. However, it is also possible, in this equation, to set the stylus radius to zero without any necessity for a simultaneous examination of the way the deformation would change with a change in that radius value.¹ If this is done, the resulting equation is of exactly the form that obtains if the groovewall waveform is replaced by a tracing-corrected one. The replacement may be effected by relabelling the left-hand side of Shiga's Eq. (20') as $q(Vt)$, while q on the right may be relabelled as $p(Vt)$ and r replaced by $-r$. In this way, Eq. (20') is reinterpreted as specifying

1. In this equation the specialization to a deformation obeying the Hertzian law has not yet been made, except that the deformation is to lie along the local normal to groovewall.

a groovewall waveform q as a tracing-corrected version of a program waveform p , and the insertion of q so specified into Eq. (20) provides a description of the reproduction of p . Thus, even when Miller factoring is not assumed, the results of the calculation may be seen to contain it. That this state of affairs may be expected to hold generally is a matter that is to be examined in more detail in a later paper [12].

Miller factoring has evidently not been widely appreciated in its exact significance, perhaps not even by Miller himself. Nevertheless, it is characteristic of his approach, and its exact significance is essentially obvious from his use of it, so that it would be fitting to identify this lasting contribution to the understanding of the stylus-groovewall interaction with his name. That its exact significance possibly eluded him may be indicated by his discussion of translation loss. This is a first-order wavelength-dependent loss appearing as a diminution in the transmission of the fundamental component at the shorter wavelengths, and attributable to deformation. Miller found that corresponding effects for the harmonic components may be neglected. Such corresponding effects were found also by Shiga [11] who chose not to neglect them. In any case, Miller appears to regard translation loss, in consequence of its approximate role, as being a separable factor in the deformation process.

P. Kantrowitz, in consultation with Miller, made the appearance of this translation-loss factor explicit by exhibiting [13] a block diagram similar to the present Fig. 1. In his diagram, Kantrowitz made the groovewall waveform subject to translation loss, the so-called G-loss function, before it is subjected to tracing-error generation.² This factoring is in contradiction to the more fundamental Miller factoring and would require various powers of the G loss to appear in the terms of corresponding order in the overall distortion. Miller found no need for such appearance, nor did Shiga. The Kantrowitz factoring is incorrect, and the Miller factoring shown here as Fig. 1 remains valid.

For the correction of tracing error, the significance of the Miller factoring shown in Fig. 1 may be understood with the help of Fig. 2. In this figure, the block diagram of Fig. 1 has been augmented by adding the block *Tracing Correction Generator*, whose action is to be conceived as forming a groovewall waveform embodying an exact complement to tracing error. The exact nature of this complement has been described elsewhere [5, 14-16]. If the groovewall waveform may be made to embody such a complement, then what had been the *Undeformed Trajectory Waveform* of Fig. 1 becomes an exact replica of the *Program Waveform* in Fig. 2, and the sole distortion process remaining is that of the *Deformation Error Generator*, including any frequency-response limitations inherent in the stylus-groovewall interaction. Since the cancellation of tracing error is thus shown, with the help of Miller factoring, to have been completed prior to the development of any frequency-response limitations in the reproducing element, it is seen

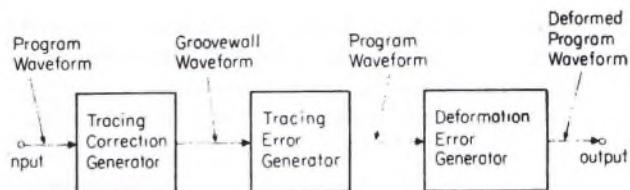


Fig. 2. The role of tracing correction in the presence of deformation.

that the efficacy of this cancellation does not depend upon any bandwidth requirement for the final reproducing element.

RELEVANCE TO THE EXPERIMENTAL EVIDENCE

To be sure, there is a low-pass loss function to be inserted between the tracing-correction and tracing-error generators. What has been shown is that this insertion is not the G-loss function of the final reproducing element nor any other aspect of that element; it is the loss function of the recording element, as shown in Fig. 3. It

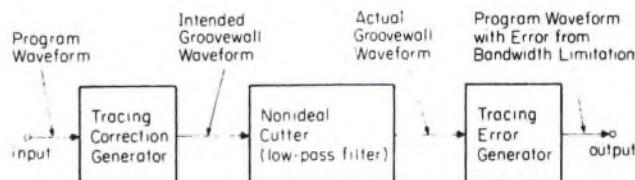


Fig. 3. The role of a bandwidth limitation in tracing-error correction.

will be necessary to inquire more deeply into the effects of this, and it will be the purpose of the section following this one to report the results of the inquiry. The effects will be seen to be rather minor and not likely to be of experimental significance for moderate recording levels.

The effects to be found, however slight, will also pertain to the intermediate, though not the final, reproducing element in the disc-to-disc dubbing method of tracing correction. Fox and Woodward in reporting their demonstration [3] of this dubbing method did show that a highly desirable, though limited, reduction in distortion was obtained. They conjectured that inadequacies in the magnitude and phase of the reproducing frequency response might have been at fault. This conjecture is not supported, because it is probably the case that the double-bandwidth requirement was met, and because a failure even in this regard, as will be seen, is not likely to have been significant. The fault more probably lies with an inadequacy of the dubbing method to provide any significant correction for deformation error, so that the compounding of that error masked a large part of the cancellation of tracing error. This compounding places the further analysis of dubbing systems beyond the scope of the present paper.

The double-bandwidth requirement was also fully met in the experimental demonstration reported by Fox and Woodward [3] of the efficacy of electronic waveshaping of the recorded waveform. In fact, the bandwidth of the recording element probably extended beyond $3f_u$ for the test frequencies used. The demonstration was made, then, with a system more closely resembling Fig. 2 than

2. Kantrowitz also represented tracing-error generation schematically as a simple transfer-curve nonlinearity. Such a representation contradicts the exact power-series representations known at the time, and it also contradicts the present-day understanding of tracing error as arising from a delay-modulation process.

Fig. 3, and essentially no test of the effect of a bandwidth limitation was made in this instance, either. It is true that the electronic waveshaping was an approximation, in another sense, to the correct shaping, but a qualitative examination [17] of this has shown it to be unlikely to be of serious consequence.

Also omitted from Fig. 2 is a representation of the role of the vertical cutting and tracking angle insofar as these would be different from the true vertical, namely 0° . A 15° angle was sought and presumably attained by Fox and Woodward both in cutting and playback. Consequently, Fig. 2 should also show the cutting-angle transformation inserted between the *Tracing-Correction Generator* and the *Tracing-Error Generator*, while the tracking-angle transformation should be shown inserted after the *Deformation-Error Generator*. As explained elsewhere [14], this insertion of the cutting-angle transformation is not at the most-effective location for avoiding a tracing-tracking interaction. Due allowance may be made for this effect in making the tracing correction, so that the interaction need not appear [16], but this was not done. Even so, the interaction may be shown to produce effects of essentially negligible consequence, in view of the fact that the demonstration was based on waveforms of rather small peak slope values.

The location of the insertion of the tracking-angle transformation after the *Deformation-Error Generator* provides the opportunity for another interaction, namely that with deformation. The tracking angle helps define the constraint by which the forces brought to bear on the groovewall are resolved. Also, with deformation intervening, the tracking-angle transformation does not appear at the theoretically-correct place in the sequence of transformations, so that there will be an interaction on more general grounds.

The general experience has been, however, that the consequences of these interactions between errors are of negligible proportions in comparison to the primary effects of these errors. On this ground, the absence from Fig. 2 of these transformations relating to cutting and tracking angles, so long as the angles are matched and not too great, is justified. It is Fig. 2, then, that is to be taken as rather accurately representing the experimental demonstration of Fox and Woodward. With the justification of this representation, it is a consequence of Miller factoring that the residual distortion of this demonstration may at last be identified for the most part. Of the long list of factors proposed by Fox and Woodward, only one emerges as being of substantial significance. The residual distortion is identified as that of deformation error.

This residual distortion has remarkable properties. In the presence of tracing error it is very large, much larger in fact than tracing-error distortion itself. On the other hand, when tracing error is not present, this residual distortion is comparable to the tracing-error distortion that would have obtained. The ability to switch the tracing error on and off in the demonstration of Fox and Woodward, essentially switching between Figs. 1 and 2, has revealed this residual distortion to interact very strongly with tracing error. So much is this the case that the interaction between these effects is seen to be stronger than either effect in pure form. Because of Miller

factoring, the structure of tracing error cannot be altered in any fashion by this interaction; the interaction is unidirectional. It is the residual distortion mechanism itself—deformation error—that is so strongly altered by the presence of tracing error, and it must be conceded, contrary to the general experience, that this interaction is scarcely a second-order phenomenon.

All of the data of Fox and Woodward may be quantitatively understood with the help of Miller factoring and the identification of the residual distortion as pure deformation-error distortion. A full discussion of these data would, however, go beyond the scope of the present paper, and must be deferred [12]. Let it suffice to present a discussion of a part of the data and only to sketch the broader implications.

A part of the data of Fox and Woodward, replotted to logarithmic scales, is shown in Fig. 4. The distortion

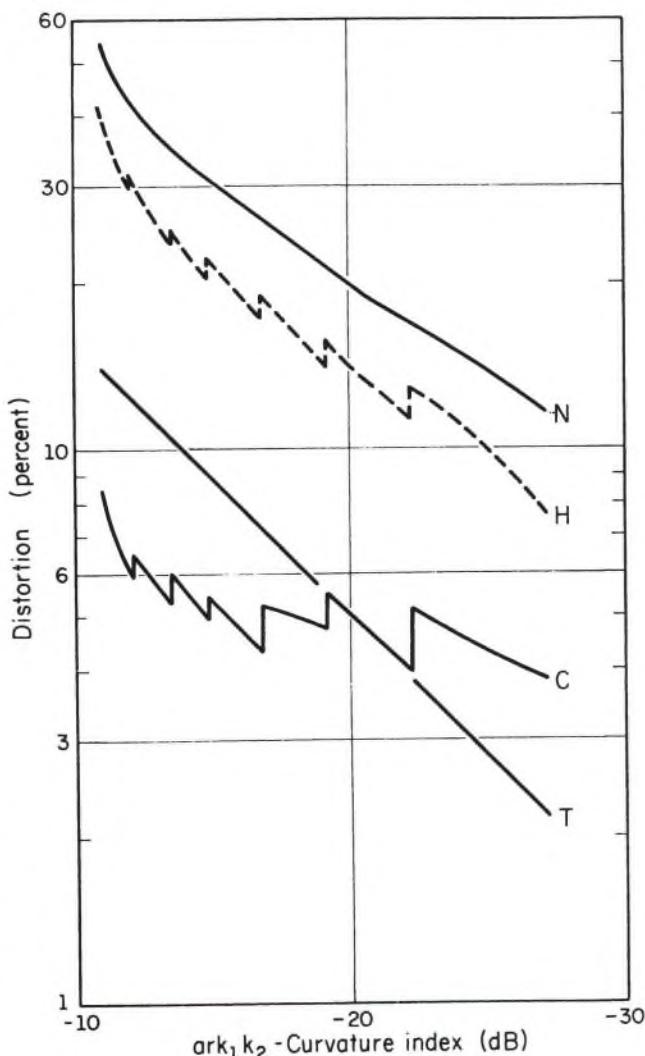


Fig. 4. Difference-tone intermodulation distortion vs curvature index. These data represent displacement measurements by Fox and Woodward replotted to logarithmic scales from Fig. 12 of Ref. [3]. **Curve N** is for the "normal" method of recording in which no allowance is made for the tracing error of a finite size for the reproducing stylus, while **Curve C** is for corrected recording for playback with an 18 μm (0.7 mil) stylus radius. The **Curve T** represents the expected tracing-error distortion in the absence of deformation error. **Curve H** represents the hypothesis that the measured improvement could be explained simply by subtracting the known amount of the correction from the "normal" distortion. The failure of this hypothesis shows that tracing correction also removes a large part of the deformation error.

percentage pertains to the difference tone at 500 Hz generated as a CCIF beat between the primary tones at 4.0 kHz and 4.5 kHz each recorded with a peak displacement of $0.89 \mu\text{m}$ (3.5×10^{-5} in.), corresponding to peak velocities of 3.16 and 3.55 cm/sec, respectively. While a vertical modulation was actually used, the specified recorded displacements are those that obtained for each groovewall separately [18], and in particular for the right-channel groovewall from which the present data were obtained. The percentage represents peak displacement of the 500 Hz beat tone as a fraction of $0.89 \mu\text{m}$.

The distortion data are plotted vs a curvature-overload index expressed in decibels, a logarithmic measure of ark_1k_2 , in which a is the wave amplitude, k_1 and k_2 are the two wave numbers (radians per unit distance) corresponding to the two frequencies recorded, and r is the stylus radius $18 \mu\text{m}$ (0.7 mil). The quantity ark_1k_2 is the peak curvature of a wave of amplitude a and frequency $\sqrt{(k_1k_2)}$. Thus the quantity ark_1k_2 in decibels represents a recorded-curvature level whose 0 dB reference is the stylus curvature. On this scale, curvature overload for a single tone of frequency $\sqrt{(k_1k_2)}$ comes at 0 dB, while for the double tone it comes at -6 dB, at which point the tracing-error beat-tone displacement would be down 12 dB from either of the primary tones, i.e., at 25%, according to theory. The straight line marked T in Fig. 4 gives a plot of the beat-tone percentages attributable to tracing error.

The overload-index scale is reversed, greater values of the index being indicated farther to the left. This was done because the variation in this index reflects variations in recording diameter, and it was thought appropriate to have the effect of the larger recording diameters appear farther to the right. The curves cover the range of recording diameters from 11.4 cm (4.5 in.) to 29.2 cm (11.5 in.). The curves marked C and H show steps marking changes in the functioning of the system at the diameters of 12.2, 13.2, 14.2, 16.0, 18.3, and 21.8 cm (4.8, 5.2, 5.6, 6.3, 7.2, and 8.6 in.).

The curve marked N gives the total difference-tone displacement for the "normal" method of recording corresponding to the conditions of Fig. 1. Since the values lie markedly above those that may be attributed to tracing distortion (Curve T), it must be understood that this mode of exciting deformation distortion makes its beat-tone measure largely additive to that of tracing error. In this, it is reasonable to assume that sufficient tracking force was applied to assure a negligible contribution from tracking-failure error, and that tracking-angle error and electronic distortions are similarly negligible. The recording was represented in vinyl pressing.

The curve marked C was obtained for a recording made in the same manner as Curve N except that RCA's Dynamic Recording Correlator (DRC) provided electronic correction of tracing error corresponding to the setup of Fig. 2. The steps arise because the DRC does not provide a correction that varies continuously with recording diameter, but provides a fixed amount within zones of the recording diameter. The boundaries of these zones mark the steps. Except for this artifact, it must be understood from Fig. 2 that the data indicate essentially pure deformation-error distortion and in a mode characterized by a nearly pure two-tone excitation, in contrast to the mode of excitation for Curve N . It will be

noticed that the recording-diameter dependence of deformation distortion is markedly different from that of tracing error.

In the absence of deformation, the difference between curves N and C would be simply explained by observing that the DRC has produced frequency components, representing the complementary distortion, that are oppositely phased to those to be generated by tracing error in the expectation of producing the desired cancellation. This is a perfectly correct point of view, though, as will be seen in the following section in which this point of view is of necessity taken, it is a complicated way of looking at things. With deformation present, however, this cancelling hypothesis appears grossly insufficient to account for the difference between Curves N and C . This hypothesis is represented by Curve H , obtained by taking the measured amounts of the complementary component produced by the DRC and subtracting these from Curve N . Clearly, the resulting values of Curve H are by no means as small as those of Curve C . Nevertheless, it must be believed that Fig. 2 truly describes the situation. It must be concluded that the magnitude of the deformation error itself has been profoundly reduced as a consequence of eliminating tracing error.

This conclusion is understandable. Near 4 kHz where the stylus impedance has a substantial inertia-reactive component, the deformation is strongly excited by those trajectory curvatures of such a sign that the burden of supplying acceleration falls upon the groovewall. These curvatures are sharper for the uncorrected trajectory than they are for those trajectories from which tracing error has been removed. Thus, the use of tracing correction has dramatically softened the excitation of deformation error in this way, and obtained a distortion reduction beyond the mere banishment of tracing error itself.

That this is not simply fortuitous may be seen from the fact that the same theme is of great power in explaining the rest of the data of Fox and Woodward, but with some variations. These cannot be dealt with here in adequate detail, but it turns out, for example, that in the sum-tone components the deformation part is largely oppositely phased, though larger, than the part attributable to tracing error. With allowance for this, the interpretation becomes quantitatively consistent with the interpretations for the difference tone. It follows of necessity that a similar situation must be expected for the harmonic tones, namely that the deformation part would be largely oppositely phased to the part arising from tracing error. This fact would explain Walton's observation [8] of a tendency of deformation error to partly cancel tracing error as being a fortuitous phasing phenomenon, one that would appear in his observation of the second harmonic, but one that would not apply to other, possibly more significant indicators of distortion.

Shiga found terms in his analysis [11] of deformation errors that are of a structure similar to that of tracing error and of a polarity to partly cancel that error. This cancelling is to be distinguished from that observed by Walton in that it would make a more consistent appearance among the various indicators, including the difference tone. Upon inspection of this term, it may be seen that it would be the more significant for the larger values of the steady part, in contrast to the fluctuating part, of the deformation, i.e., for the larger tracking forces or

the sharper styli. In the data of Fox and Woodward, the part for the left groovewall pertains to the larger tracking force because of the side thrust or skating force. In these data, the appearance of Shiga's term may be clearly seen, though the measured values of the uncorrected distortion were never less than the distortion to be expected from tracing error. For this term to be significant, however, it is necessary for the steady part of the deformation to be a substantial fraction of the stylus radius, a condition very likely to leave permanent deformation indicative of its destructive magnitude. It is difficult to maintain that this mode of obtaining partial relief from tracing error should be deliberately sought in high-quality systems.

The mechanical impedance behind the 18 μm stylus used in the playback part of the demonstrations of Fox and Woodward was probably not as low as would be possible in today's state of the art, so that the deformation contributions, including the interaction excited by tracing error, are perhaps atypical of results that might be currently measured, at least for an 18 μm stylus. In more recent difference-tone measurements [19], unfortunately not made both with and without a facility for cancelling tracing error, there appears to be some reduction in total distortion with more nearly modern styli, though not as much as might have been hoped. The reason is probably that the seriousness of the deformation, while reduced by virtue of using a lower mechanical impedance, was aggravated by using the sharper curvature of the elliptical tip, possibly resulting in an overall worsening of the deformation. At the same time, of course, the tracing error was reduced somewhat. If the data in the new measurements are to be compared with the old, allowance must be made for the fact that the newer percentages are taken on a velocity scale in which the base is the sum of the peak velocities of the test tones, resulting on both counts, in very small percentages. At 1000 Hz, for example, these newer percentages must be multiplied by a factor of 5 to convert to the older displacement basis, whereas at 10 kHz the appropriate factor would be 50.

It would be interesting to see the older measurements repeated using today's reduced mechanical impedances, to see what progress has actually been possible against deformation. Recently-made groovewall photographs [20] do show the kinds of permanent damage that may result from deformation, but they do not give the scale of the transient deformation. So far, waveform distortion is the only way to measure that. It may be argued, then, that since it is now understood that deformation distortion may be measured in a way essentially free of contamination by other errors, such measurements should proceed. Since these same photographs do show the elliptical stylus to be more damaging than the 18 μm round stylus, when compared at the same mechanical impedance, it should be especially interesting to see what new lows in overall distortion are now attainable with tracing correction where no risk of the elliptical damage need be incurred.

RELEVANCE TO THE RECORDING BANDWIDTH

Thus far, it has been possible to show that the

bandwidth requirement for achieving the desired cancellation of tracing-error distortion is inapplicable to the final reproducing element, because of Miller factoring. Also, while the experimental demonstrations of Fox and Woodward [3] showing a very desirable cancellation did not challenge the applicability of the bandwidth requirement, Miller factoring was found to be very helpful in sketching the kinds of interpretations that could be drawn. It is still necessary to examine this requirement more deeply to be sure that the applicable requirement should not actually be more strictly drawn, since perfect cancellation would seem to require an infinite bandwidth. Thus it is possible, in principle, for a significant part of the residual distortion measured by Woodward and Fox to be attributable to a failure to obey a stricter bandwidth requirement. On the other hand, it is possible that a much weaker requirement would be acceptable as a practical matter.

The question may be largely settled by making trial calculations for the system of Fig. 3. Since the frequency response of the final reproducing element and, indeed, all groovewall-deformation phenomena are irrelevant to the question, these aspects have been omitted from the calculations and do not appear in Fig. 3. The calculations are rather complicated, so that instead of seeking to describe them by means of a single formula, it was thought better to discuss numerical examples and leave most of the mathematical details for an appendix.

For the numerical example, it was decided to choose two tones placed near the upper edge of the audio spectrum, and to choose the tones to be of equal displacement amplitude. Their sum is to represent a program waveform of such an amplitude as to just attain curvature overload in one case, and in another case to fall 6 dB short of overload. On an arbitrary frequency scale, these tones were chosen to have the frequencies 9 and 10, and it was decided to regard the audio spectrum as extending to a frequency just above 10, while the double-bandwidth spectrum would extend to a frequency just above 20.

The tracing-error distortion products for this two-tone signal were then computed using the methods indicated in the Appendix, with the results shown in the third column of Table I. The displacement amplitudes shown there are based on the value 100 for the primary tones, as indicated in the second column. The values given in parentheses are for the 6 dB below overload case. It may be seen that in addition to the usual first-order combination tones deriving from the second-order part of the nonlinearity, namely first-order difference and sum frequencies together with second-harmonic tones, there are some second-order combinations, $8 = 2 \times 9 - 10$ and $11 = 2 \times 10 - 9$, though these last are rather weak.

The presence of the tone at a frequency of 11 caused a revision of the choice of cutoff frequency. Clearly, the exclusion or inclusion of such a weak tone was not going to make for a substantial difference in the calculated result. On the other hand, it did not seem reasonable to assume that the filter should be modeled as having such a sharp cutoff as to include 10 but exclude 11, especially since it was desired to let the model be one having a linear phase response. Since there was more "room" between 11 and 18, it was finally decided that the model would be less vulnerable to criticism if the cutoff were

TABLE I. Numerical examples of an interaction of tracing correction with a bandwidth limitation. In the second through fifth columns are tabulated the values of the frequency components relative to 100 on a displacement basis, and for a recording level verging on curvature overload, except for the values in parentheses which are for a level 6 dB below overload. The values in the third column are also the ones which would obtain in the output if no correction were made. The output is defined as that for Fig. 3, in that no accounting for the subsequent effect of deformation-error generation is made.

Frequency	Program Waveform	Intended Groovewall Waveform	Output Waveform Cutoff > 11	Cutoff > 20
1	00.0	-25.0 (-12.5)	-5.8 (-0.6)	-2.3
2	00.0	00.0	-1.3	-0.7
8	00.0	-2.4	-4.1	1.0
9	100.0	91.5 (97.8)	84.5	94.7
10	100.0	90.3 (97.5)	83.3	93.9
11	00.0	-3.6	-5.2	-1.6
18	00.0	8.0	00.0	-1.9
19	00.0	19.3	00.0	3.2
20	00.0	9.6	00.0	-1.5

placed in that region, for the "single-bandwidth" case. For the "double-bandwidth" case there would be plenty of room between the frequencies 20 and 27 to place that cutoff. In the table, the effects are labeled as pertaining to cutoffs above 11 and above 20 respectively.

For the single-bandwidth case, the greatest distortion component is the one at the frequency 1, and it is the one for which the most nearly complete cancellation would be desired. Principally, this is to be accomplished with the beat between the tones at 9 and 10 of strengths 91.5 and 90.3, now replacing their original strengths. Because of the reduction in strength, however, the cancellation cannot be complete on that score alone. Some help (but not much) is provided by the fact that the beat to be cancelled is not transmitted through the tracing error generator with quite its full amplitude. Worse than that, the beat between 8 and 9 and between 9 and 11 also contributes, but the sign is wrong for cancellation.

For the double-bandwidth case, additional cancellation is available from the beats between 18 and 19 and between 19 and 20, so that a further reduction in amplitude at the frequency 1 is obtained. Generally, the availability of additional ways of generating combination tones makes for a more nearly complete cancellation of the spurious tones within the original band, just when these combinations may be drawn from the above-band spurious components. By exploring these combinations in detail, it is seen that every distortion component in the single-bandwidth case has been reduced further in the double-bandwidth case, while the strengths of the original tones have been partly restored. Also, of course, the above-band harmonics and sum tones were reduced.

While it is seen that providing a doubling of bandwidth for the recording element results in a substantial improvement in the desired cancellation, it is also seen that this doubling is not sufficient to provide perfect cancellation, and that one could expect further improve-

ment upon a tripling or quadrupling of bandwidth, since perfect cancellation is available only if there are no limitations in bandwidths. On the other hand, it is seen that a very desirable cancellation is obtained without any doubling of bandwidth, especially at curvature levels below overload.

The below-overload cancellation is indicated by the figures shown in parentheses in Table I. These were computed for peak waveform curvatures half as sharp as the stylus curvature, i.e., at 6 dB below overload. As before, the reference displacement amplitude is taken to be 100. These primary tones at the frequencies 9 and 10 are transmitted with only very slightly diminished amplitudes so that there is very little diminution in the beat-tone strength to be used in cancelling the original beat tone. For simplicity, the full spectrum of distortion components were not calculated in this case. The particular beat-tone cancellation shown is sufficient to indicate the extreme sensitivity to curvature level in the face of a bandwidth limitation. From this indication, it may be estimated that reducing the curvature level from overload to one some 2 dB below overload would be as effective as doubling the bandwidth.

It should be noted that the demonstrations of Fox and Woodward [3], were for curvature levels 5 dB and more (-11 dB or lower on the index scale used for Fig. 4) below curvature overload. More than that, the available bandwidth was ample with respect to the hypothetical double-bandwidth requirement, and even the linear-phase (constancy of group delay) hypothesis was probably reasonably well satisfied over the limited band actually used. It is clear, then, that no effect of a bandwidth limitation could be said to have contributed in a major observable way to their results. While an extension of bandwidth would not be required to accommodate tracing correction for handling program material of a bandwidth near that of the cutter, it would appear appropriate for the level restriction to be obeyed, and for a linear-phase equalization of the cutter to be employed over as wide a part of its band as possible.

DISCUSSION

While it is evident that there exists an interaction between tracing correction and a bandwidth limitation, it has been shown that this interaction is very weak. Because it is so weak, a very modest limitation in recording level, of the order of -2 dB with reference to the curvature-overload level, is as effective as a doubling of bandwidth in maintaining a specified high order of tracing-distortion cancellation. Cancellation of tracing error that is essentially complete, for practical purposes, is thus obtainable within bandwidths no wider than needed to pass the program signal without regard for above-band components generated in the tracing correction. Any above-band components in the program signal ought also, of course, to be filtered out prior to making the correction.

With the help of Miller factoring, moreover, it has been shown that groovewall deformation processes play no role in either the generation or the correction of tracing error if the latter is done by synthetic means of waveform shaping. Thus, full correction of tracing error predictable by theory should be available in practice.

Essentially the whole of the residual distortion appearing in the tracing-correction experiments of Fox and Woodward has been shown, again with the help of Miller factoring, to be deformation-error distortion. In these experiments, it has made its appearance in pure form for the first time. This identification has also provided a means of accounting for the very large difference-tone distortion figures in comparison with the sum-tone figures in those instances in which the tracing distortion was not cancelled, as being in large part a phasing phenomenon. The deformation distortion, in pure form comparable to tracing distortion, interacts with the latter in a way to greatly aggravate the deformation; this aggravated distortion is then augmented by tracing distortion for the difference tone, but diminished by it for the sum tone. Tracing correction removes these very large interaction contributions along with removing tracing error itself, leaving only the unaggravated or pure deformation-error distortion.

If the Fox-and-Woodward experiments were to be repeated with modern styli backed by a lower mechanical impedance, it might be expected that, instead of the aggravated deformation distortion being several-fold more serious than tracing-error distortion, it might be comparable to it, while the unaggravated or pure deformation distortion, the residual distortion, would be substantially less than the tracing-error distortion. This situation may be compared with the efficacy of an alternative approach to the reduction of tracing error, the use of the so-called elliptical stylus, for example, one with contact radii of curvature of $6 \times 18 \mu\text{m}$, instead of the spherical $18 \times 18 \mu\text{m}$. If these be compared at identical mechanical impedance and tracking force, it is clear that the tracing error will be reduced for the $6 \times 18 \mu\text{m}$ stylus, but at the expense of an increase in deformation error. The tracing error will have been reduced by a factor 3 (not cancelled); the interaction with deformation will have been reduced on the one hand, because of this reduction in tracing error, but increased, on the other, because of the increase in deformation. Also, the "pure" contribution from deformation will itself be greater. The net result should be that no significant overall reduction in distortion may be attributed to the use of an elliptical stylus in comparison to that obtaining from the use of a spherical one of the same major radius and backed by the same mechanical impedance.

Preliminary listening tests, when the comparison may be made at identical mechanical impedances, tend to support this conclusion. Earlier tests, however, had seemed to indicate the contrary; it is believed that the fact that the commercial practice had been to fit elliptical styli only to assemblies with improved impedance characteristics had tended to mask the obviously greater deformation potential of the elliptical stylus. That the elliptical stylus has the greater deformation potential should be obvious from a comparison of the radii, $6 \times 18 \mu\text{m}$ vs $18 \times 18 \mu\text{m}$. The "ocular proof," however, is available in the remarkable groovewall photographs made with a scanning electron microscope by Woodward, Coutts, and Levin [20]. It is to be hoped that these photographs will motivate the commercial offering of $18 \times 18 \mu\text{m}$ styli fitted to state-of-the-art stylus assemblies, a fitting hitherto thought "uncommercial."

It is further to be hoped that additional measurements

of the Fox-and-Woodward type will be made with state-of-the-art stylus assemblies, since photographs of the remanent deflection are incapable of showing the larger active deformation error. The identification of the latter, in its first appearance in pure form, as the residual distortion in tracing-corrected recording should greatly facilitate further studies of deformation distortion. For example, with the use of tracing correction, the sole remaining distortion mechanism, deformation-error distortion, may be excited by waveforms of a prescribed simple frequency structure, so that the interpretation of the measured results is greatly simplified.

The ability to make theoretical predictions of deformation-error distortion should also be facilitated if no allowance need be made for tracing error. The theoretical formulations of Shiga [11] have been shown to be substantially simplified if specialized to the tracing-corrected case. With the help of this simplification it is to be expected that analyses for more realistic deformation laws than the Hertzian one may be developed soon.

With the advance of the date that more realistic analyses may be made, there should be an advance of the date at which more meaningful comparisons between theory and experiment may be made. Thus, one may predict an overall advance of the general understanding of this last significant distortion-generating mechanism. With this will come an advance in means of control, leading to substantial reductions in overall distortion.

Some means of control are already at hand. With tracing error eliminated, it is clear that, at every stage of the development of the art, the slighter deformation error will always be incurred with the more gently curved stylus tip backed by the smaller mechanical impedances. Both these means of control will always be available for commercial exploitation, but only if tracing error is eliminated by means independent of them. On the contrary side, the choice of more sharply curved styli for the playing of uncorrected recordings is seen to be an obsolescent practice actually inhibitory to any future development of the art.

At present, the largest stylus radius that is compatible with current standards is $18 \mu\text{m}$ (0.7 in.). It would appear from the above argument that the future development of the art is most strongly favored, in a way compatible with immediate realizations of substantial improvement, if this largest compatible stylus radius is chosen and the manufacture and sale of all others discouraged. The rewards of tracing correction, the elimination of tracing error combined with a substantial reduction in deformation error, will thus provide the necessary ground for the future development of the art, a development not available by any other route.

APPENDIX

Multitone Distortion Analysis

The parametric equations describing the tracing-error transformation of the slope waveform $\sigma(\zeta)$ to produce the distorted slope waveform $s(z)$ are of the form of a skew-coordinate transformation [14, 15], namely

$$s(z) = \sigma(\zeta), \quad (\text{A1a})$$

$$z = \zeta - rg(\sigma), \quad (\text{A1b})$$

in which r is the zero-slope radius of curvature of the

stylus profile and $g(\sigma)$ describes the shape of the profile by specifying the abscissae δ_z along the profile as a function of the slope values σ . In this form, the equation of the profile is

$$\delta_z = rg(\sigma). \quad (\text{A1c})$$

This equation is sufficient to specify the profile; in terms of it, the ordinates δ_y along the profile are

$$\delta_y = rg(\sigma)\sigma - r\int g(\sigma)d\sigma. \quad (\text{A1d})$$

In these equations, z and ζ are the abscissae along the quiescent-groove axis.

The Fourier integral for the spectrum $S(k)$ of $s(z)$, in which k is the space frequency (radian wave number), is

$$S(k) = \int e^{-jkr} s(z) dz = (1/jk) \int e^{-jkz} s'(z) dz, \quad (\text{A2})$$

in which the integrals are definite ones covering $-\infty < z < \infty$. The second form of the integral of Eq. (A2) may be expressed in terms of the quantities of Eqs. (A1) as

$$S(k) = (1/jk) \int e^{-jk\xi} e^{jkrg(\sigma)} \sigma'(\xi) d\xi, \quad (\text{A3a})$$

in which integration by parts may be performed with the result

$$S(k) = \int_{-\infty}^{\infty} e^{-jk\xi} \left[\int_0^{\sigma(\xi)} e^{jkrg(\tau)} d\tau \right] d\xi \quad (\text{A3b})$$

in which the limits of integration have been explicitly indicated. For the case of the linear skew transformation, $g(\tau) = \tau$, the case of the parabolic stylus profile, the integration of parts may be explicitly completed, so that Eq. (A3b) may be written as

$$S(k) = (1/jkr) \int e^{-jk\xi} [e^{jkrs(\xi)} - 1] d\xi. \quad (\text{A3c})$$

Equations (A3) are valid whenever both $s(z)$ and $\sigma(\xi)$ satisfy the hypotheses of the Fourier-integral theorem, including the requirement that these be single-valued functions.

The single-valued requirement is worthy of explicit notice, because if either of $|rg'\sigma'|$ or $|rg's'|$ be allowed to exceed unity, then even if one of s,σ is single valued the other will not be. This requirement makes it impossible to consider an extension of the class to which s,σ belong so that it might include such generalized functions [21] as the Dirac delta, for example. On the other hand, there are no grounds for excluding generalized functions from the class to which $S(k)$ belongs, and, in particular, it is permissible for $s(z)$ and $\sigma(\xi)$ to be periodic. In the periodic case, the interpretation of $S(k)$ as a generalized function may be avoided in the usual manner by restricting the range of integration to span one period λ and performing the evaluation only for values of k that are integer multiples, $k_n = nk_1 = 2\pi n/\lambda$, of the fundamental frequency k_1 . Then, $S_n = S(k_n)/\lambda$ is the ordinary Fourier-series coefficient for which

$$s(z) = \sum S_n \exp(jk_n z) \quad (\text{A4})$$

provides for the elimination of the parameter ξ from Eqs. (A1).

The evaluation of the integral (A3c) is known to be a tractable problem in a few instances, one of which obtains for $\sigma(\xi) = s_1 \sin k_1 \xi$. In that instance, the integral is

$$S_n = (1/2\pi jr) \int e^{-jk\xi} [\exp(jrks_1 \sin k_1 \xi) - 1] d\xi, \quad (\text{A5a})$$

in which $k = k_n = nk_1$ and the range of integration spans one period. The result is that $S_n = 0$ and

$$S_n = J_n(nr k_1 s_1) / nr k_1, \quad (\text{A5b})$$

as first obtained by Bessel [22] in terms of his famous function of the first kind, order n , $J_n(x)$. The corresponding trigonometric series is just the sine series with coefficients

$$2jS_n = 2J_n(nr k_1 s_1) / nr k_1. \quad (\text{A5c})$$

For $\sigma(\xi) = s_1 \sin k_1 \xi + s_2 \sin k_2 \xi$, the analogous result may be obtained by noting that

$$e^{jkrs(\xi)} = \exp(jkr s_1 \sin k_1 \xi) \sum_n J_n(kr s_2) \exp(jnk_2 \xi), \quad (\text{A6})$$

so that the sine-series coefficients are

$$2 \sum J_m(rs_1 k_{mn}) J_n(rs_2 k_{mn}) / rk_{mn} \quad (\text{A7a})$$

for the frequency

$$k_{mn} = mk_1 + nk_2, \quad (\text{A7b})$$

where the sum is to range over all positive and negative values of m and n such that at least one of these is different from zero, and such that k_{mn} is maintained at a fixed positive value. In principle, if k_1 and k_2 are commensurate, there are infinitely many terms of the same frequency k_{mn} . For example, the value k_{01} , which is simply k_2 , can be obtained for $m = 0, n = 1$ by definition. However, for $9k_2 = 10k_1$, the same value is obtained for $m = -10, n = 10$, and also for $m = 10, n = -8$, etc. Fortunately, the properties of the Bessel functions are such that if the single-valued conditions be met the higher-order combinations will make very small, usually negligible, contributions, with the result that as a practical matter each Fourier-series coefficient is obtained as a single term.

In using these results for the calculation of the third column of Table I, integration was used to convert the formulae to the description of displacement waveforms of amplitude a_1 and a_2 with $s_1 = a_1 k_1$ and $s_2 = a_2 k_2$. The coefficient corresponding to Eq. (A7a) then is

$$2J_m(ra_1 k_1 k_{mn}) J_n(ra_2 k_2 k_{mn}) / rk_{mn}^2. \quad (\text{A8a})$$

As a fraction of the mean peak displacement of the primary tones, $a_{12} = \frac{1}{2}(a_1 + a_2)$, this is

$$2J_m(ra_1 k_1 k_{mn}) J_n(ra_2 k_2 k_{mn}) / ra_{12} k_{mn}^2. \quad (\text{A8b})$$

In the CC1F-1M test, one would have $a_1 = a_2 = a$, in which case Eq. (A8b) becomes

$$2J_m(rak_1 k_{mn}) J_n(rak_2 k_{mn}) / rak_{mn}^2, \quad (\text{A9a})$$

which upon specialization to the first-order difference tone, for example, is

$$-J_1(rak_1 \Delta) J_1(rak_2 \Delta) / rak^2, \quad (\text{A9b})$$

for $\Delta = k_2 - k_1$. In this, one may use the expansion

$$J_n(x) = (x/2)^n [1 - (x/2)^2/(n+1) + \dots] / n! \quad (\text{A10})$$

to evaluate Eq. (A9b) as

$$- \frac{1}{2} rak_1 k_2 \quad (\text{A9c})$$

to within a very small error. Curvature overload, the

verging on the formation of a multiple-valued waveform, obtains for

$$ra(k_1^2 + k_2^2) = 1, \quad (\text{A11a})$$

but since

$$k_1^2 + k_2^2 = (k_2 - k_1)^2 + 2k_1 k_2$$

it may be seen that when k_1 and k_2 differ by a small percentage of their own values

$$\frac{1}{2}(k_1^2 + k_2^2) \approx k_1 k_2$$

so that Eq. (A11a) may be approximately replaced by

$$rak_1 k_2 = \frac{1}{2} \quad (\text{A11b})$$

to specify overload. Thus, from Eq. (A9c) it may be seen that at overload the first-order difference tone is $\frac{1}{4}$ of the displacement a , a distortion percentage of 25%. The sum-tone percentage is calculated in the same way, except that $k_2 - k_1$ is replaced by $k_2 + k_1$ and that there is no minus sign as appears in Eq. (A9b). This sign is a consequence of the rule

$$J_n(-x) = J_{-n}(x) = (-1)^n J_n(x). \quad (\text{A12})$$

Also, $k_2 + k_1$ is greater than $k_2 - k_1$, so that many more terms of the series Eq. (A10) must be used. For $k_1 = 9$, $k_2 = 10$ and an ra value obeying Eq. (A11b), the sum-tone percentage may be calculated as 19.31%, using tables.

In calculating the fourth and fifth columns of Table I, the interpretation was made that all of the tones appearing in the third column were primary tones and that the sign of r had been reversed. In such a case, the procedure is rather complicated and equally complicated to describe. In order to give a sketch of the procedure, let it be assumed that the number of primary tones is 5, i.e., that the signal has the five Fourier components a_1, a_2, a_3, a_4, a_5 , with frequencies k_1, k_2, k_3, k_4 , and k_5 . (The actual values of these frequencies, in one case, were 1, 8, 9, 10, and 11.) The output Fourier-series coefficients are given by sums of terms of the form

$$2J_l J_m J_p J_q / rak_{lmnpq}^2, \quad (\text{A13a})$$

as fractions of a , in which the argument of each Bessel function is

$$ra_y k_y (lk_1 + mk_2 + nk_3 + pk_4 + qk_5) \quad (\text{A13b})$$

where $y = 1$ in J_l , $y = 2$ in J_m , $y = 3$ in J_n , $y = 4$ in J_p , and $y = 5$ in J_q . For a fixed value of

$$k_{lmnpq} = lk_1 + mk_2 + nk_3 + pk_4 + qk_5 \quad (\text{A13c})$$

it may be seen that there are several combinations of the integers l, m, n, p , and q that satisfy Eq. (A13c). However, as a practical matter, it is found that the product Eq. (A13a) is negligibly small for the ra_y values to be used if any of these integers is greater than 2 and if fewer than three of these are different from 0. These approximation criteria reduced the number of significant combinations to a manageable few. Nevertheless, Eq. (A13a) has to be evaluated for each of these combinations obtaining for a fixed value of k_{lmnpq} , and the results summed with due regard for sign invoking Eq. (A12), a tedious job.

For signals well below overload, i.e., for

$$rak_1 k_2 \ll \frac{1}{2} \quad (\text{A14})$$

and $k_1 \approx k_2$, the primary-tone transmission,

$$2J_1(rak_1 k_2) J_0(rak_1 k_2) / rak_1^2 \quad (\text{A15a})$$

for one tone and

$$2J_1(rak_1 k_2) J_0(rak_1^2) / rak_2^2 \quad (\text{A15b})$$

for the other, becomes approximately

$$[1 - (\frac{1}{8})(rak_1 k_2)^2] [1 - (\frac{1}{4})(rak_1 k_2)^2], \quad (\text{A15c})$$

which may be approximated further as

$$1 - (\frac{3}{8})(rak_1 k_2)^2. \quad (\text{A15d})$$

Under these conditions of approximation, the difference tone arising from these primary-tone transmissions is

$$\frac{1}{2} rak_1 k_2 [1 - (\frac{3}{8})(rak_1 k_2)^2], \quad (\text{A16a})$$

and, in the cancellation of tracing error, this tone is to cancel the original difference tone

$$\frac{1}{2} rak_1 k_2, \quad (\text{A16b})$$

transmitted essentially undiminished. The cancellation leaves a residue that is approximately

$$(\frac{3}{8})(rak_1 k_2)^3 = (\frac{3}{64})(2rak_1 k_2)^3, \quad (\text{A17})$$

exhibiting the third-order dependence upon level.

Equation (A17) is not very accurate for curvature levels approaching overload, in violation of Eq. (A14), as is to be expected; the value $3/64 = 4.7\%$ does not compare too well with the value 5.8% shown in the third column of Table I. However, at 6 dB below overload, $2rak_1 k_2 = \frac{1}{2}$, so that the value $3/512 = 0.59\%$ is obtained, in good agreement with the tabular value. If Eq. (A17) is to yield the value 2.3%, then there must obtain $2rak_1 k_2 = 0.4$ ($23/3\% = 0.79$), a value representing a curvature level 2 dB below overload.

REFERENCES

1. W. MacNair, discussion on J. A. Pierce and F. V. Hunt, "Distortion in Sound Reproduction from Phonograph Records," *J. Soc. Motion Picture Engrs.* **21**, 157 (Aug. 1938).
2. F. V. Hunt, "The Rational Design of Phonograph Pickups," *J. Audio Eng. Soc.* **10**, 274 (1962).
3. E. C. Fox and J. G. Woodward, "Tracing Distortion—Its Cause and Correction in Stereodisk Recording Systems," *J. Audio Eng. Soc.* **11**, 294 (1963).
4. H. Redlich and H. J. Klemp, "A New Method of Disc Recording for Reproduction with Reduced Distortion: the Tracing Simulator," *J. Audio Eng. Soc.* **13**, 111 (1965).
5. D. H. Cooper, "Continuous Delay Regulator for Controlling Recording Errors," *J. Audio Eng. Soc.* **14**, 7 (1966), and *IEEE Trans. on Audio and Electroacoustics AU-14*, 16 (March 1966).
6. M. Ohkawa, "Reduction of Tracing Distortion in Disk Records by Taylor Expansion," *J. Acoust. Soc. Japan* **23**, 62 (March 1967).
7. F. G. Miller, "Stylus-Groove Relationships in Phonograph Records," doctoral dissertation, Harvard University (Issued as *Tech. Memo No. 20*—now out of print—Acoustics Res. Lab., Harvard Univ.; March 1950).
8. J. Walton, "Stylus Mass and Distortion," *Wireless World* **69**, 171 (April 1963).

9. W. D. Lewis and F. V. Hunt, "A Theory of Tracing Distortion in Sound Reproduction from Phonograph Records," *J. Acoust. Soc. Am.* **12**, 348 (Jan. 1941).
10. H. Hertz, *Miscellaneous Papers* (MacMillan and Co. Ltd., London, 1896), pp. 146-183. See also S. Timoshenko and J. N. Goodier, *Theory of Elasticity* (McGraw-Hill Book Co., New York, 1951), 2nd ed., pp. 372-377.
11. T. Shiga, "Deformation Distortion in Disc Records," *J. Audio Eng. Soc.* **14**, 208 (1966).
12. D. H. Cooper, "Interaction of Tracing and Deformation Errors" (in preparation).
13. P. Kantrowitz, "High-Frequency Stylus-Groove Relationships in Phonograph Cartridge Transducers," *J. Audio Eng. Soc.* **11**, 250 (1963).
14. D. H. Cooper, "Unified Analysis of Tracing and Tracking Error," *IEEE Trans. on Audio AU-11*, 141 (July-Aug. 1963).
15. D. H. Cooper, Integrated Treatment of Tracing and Tracking Error," *J. Audio Eng. Soc.* **12**, 2 (1964).
16. D. H. Cooper, "Interaction of Tracing and Tracking Error," *J. Audio Eng. Soc.* **13**, 138 (1965).
17. D. H. Cooper, "Construction of Correlator Waveforms," *J. Audio Eng. Soc.* **13**, 236 (1965).
18. J. G. Woodward, private communication.
19. J. G. Woodward and R. E. Werner, "High-Frequency Intermodulation Testing of Stereo Phonograph Pickups," *J. Audio Eng. Soc.* **15**, 130 (1967) referring to test record 12-5-105.
20. J. G. Woodward, M. D. Coutts and E. R. Levin, "The Scanning Electron Microscope—a New Tool in Disc-Recording Research," *J. Audio Eng. Soc.* **16**, 258 (July 1968).
21. M. J. Lighthill, *Introduction to Fourier Analysis and Generalized Functions* (Cambridge Univ. Press, London, 1958).
22. G. N. Watson, *Theory of Bessel Functions* (The MacMillan Co., London, 1944), pp. 552, 553.

THE AUTHOR



Duane H. Cooper was born in Gibson City, Illinois in 1923. He received the B.S. and Ph.D. degrees in physics from the California Institute of Technology, Pasadena in 1950 and 1955, respectively.

Since 1954 he has been with the Coordinated Science Laboratory (then Control Systems Laboratory) of the University of Illinois, Urbana, where he is presently Associate Professor of Electrical Engineering and Physics. His research interests involve statistical and information theoretic problems in instrumentation and information transmission systems with applications to the use of digital computers. In his consulting work,

he has published a number of analyses of distortion in disc-recording systems. Dr. Cooper teaches courses in circuit analysis and in communication systems at the Laboratory.

Dr. Cooper has served on the Editorial Board of the Review of Scientific Instruments and serves as referee for a number of professional journals. He is a member of the Acoustical Society of America, a member of the American Physical Society, a Senior Member of the Institute of Electrical and Electronics Engineers, and a Fellow of the Audio Engineering Society, which presented him with its Emile Berliner Award in 1968.

Development and Application of a New Tracing Simulator*

DIETER BRASCHOSS

Georg Neumann GmbH Electroacoustic, 1 Berlin 61, Germany

The playback tracing distortion, which is particularly pronounced in the reproduction of stereophonic discs, can be partially compensated by introducing predistortion into the cutterhead driving channel during mastering. An all solid-state device designed to fulfill this requirement is described.

I. GEOMETRIC DISTORTION IN DISC REPRODUCTION

Tracing Distortion

Recording on disc is accomplished by means of a sharp-edged pointed stylus. Neglecting lacquer spring back, the shape of the groove exactly follows the motion of the stylus. Figure 1 shows the cross section of a vertically modulated groove. While the recording is made with a

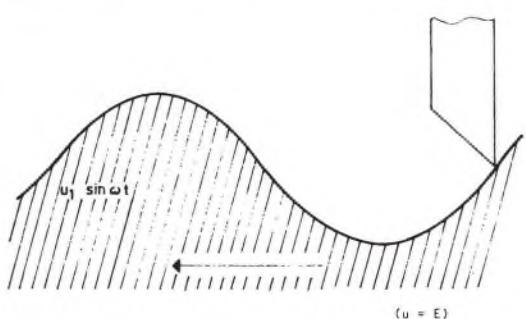


Fig. 1. Recording.

* Translated and presented by Stephen F. Temmer, Gotham Audio Corporation, New York, N.Y., on May 4, 1970, at the 38th Convention of the Audio Engineering Society, Los Angeles.

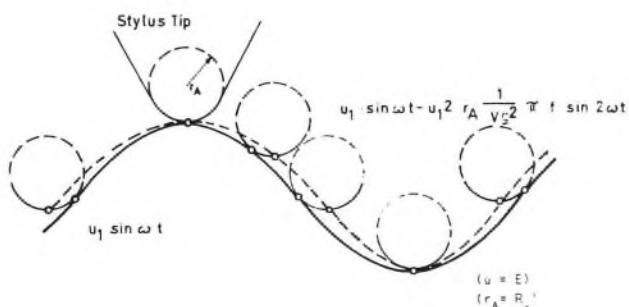


Fig. 2. Reproduction.

sharp-edged stylus, playback is by means of a spherically rounded cone. Figure 2 shows clearly that the motion of the playback stylus, which is the determining factor for the resulting signal voltage, no longer corresponds to the groove's shape. The point of contact between the spherical playback stylus and the groove wall wanders, causing the stylus to trace a path indicated by the dashed line. As a result, the sine-wave shape of the recording is reproduced in distorted form. It is fairly easy to determine from the shape of the traced curve that the distortion is mainly a matter of the second harmonic. When playing this groove back with a velocity-sensitive pickup, we obtain besides the fundamental $E_1 \sin \omega t$ the second har-

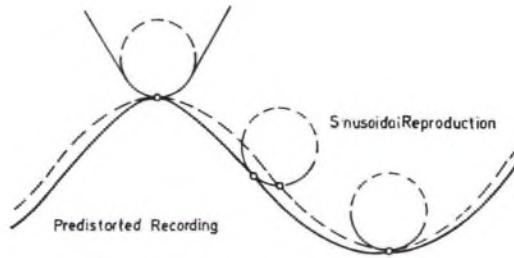


Fig. 3. Predistortion during cutting.

monic of magnitude

$$E_2 = -E_1^2 \cdot R_s (1/V_g^2) \pi f \sin 2\omega t \quad (1)$$

where

R_s = stylus tip radius

V_g = groove velocity

f = frequency.

This equation may be qualitatively interpreted by referring to Fig. 2. The higher the recorded level, the steeper the groove tracing, and the greater the deviation of the stylus-groove contact point from the stylus center line. The larger the stylus tip radius, the greater is also this deviation. If the playback stylus tip were pointed, i.e., the tip radius is zero, there would be no need for corrective action. The greater the groove velocity, the larger the recorded wavelength. Since the curve thereby becomes flatter, it also reduces the discrepancy between stylus-groove contact point and stylus center line. This means that the distortion to be compensated increases with decreasing record diameter, since the groove velocity diminishes toward the record center. The higher the frequency, the smaller the wavelength, the steeper the curve shape. This is why the distortion increases with rising frequency. If playback is to be distortion free, the recording must be made in a predistorted manner (Fig. 3). This means that the basic signal $E_1 \sin \omega t$ must have added to it a proportion of the second harmonic as first set forth in Eq. (1), however, with opposite polarity:

$$E_{2k} = +E_1^2 \cdot R_s (1/V_g^2) \pi f \sin 2\omega t. \quad (2)$$

The preceding observations are valid only for the case where both recording and playback result from only a single modulated groove flank. For a stereophonic recording the equations pertain to only a single channel. For vertical modulation the distortion products are geometrically additive, the same as the modulation signals.

Distortion as a Result of the Pinch Effect

Aside from the tracing distortion which appears in the

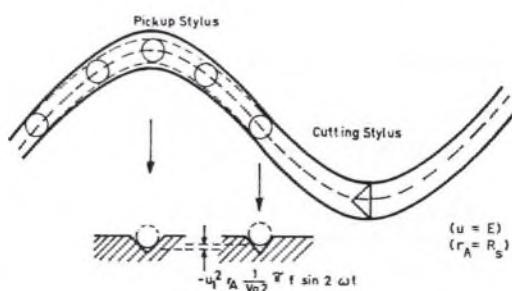


Fig. 4. Pinch effect.

vertical component, there appears distortion in the lateral direction as a result of the so-called pinch effect. Since the recording is made with a triangularly shaped stylus, lateral modulation results in a groove whose width measured at right angles to the base line of cut, does not remain constant (Fig. 4). At the point of the zero crossing (A, A') it is narrower than at its maximum width (B, B'). This produces the result that during the playback the stylus sinks deeper into the groove at point B than at point A , causing a vertical component to be added to the lateral playback, whose principal component is second harmonic.

With a modulation signal in each channel of $E_1 \sin \omega t$, the distortion magnitude is equal to $E_2 = -E_1^2 \cdot R_s (1/V_g^2) \pi f \sin 2\omega t$. In other words, the pinch effect distortion is identical to the tracing distortion (1) and also appears in the vertical direction. This means that both tracing distortion and pinch effect distortion may be compensated for a single channel by application of only one compensating signal, Eq. (2).

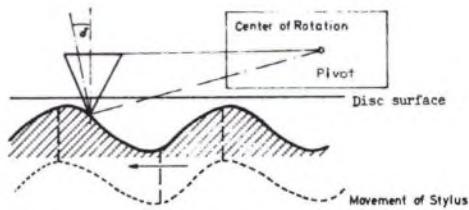


Fig. 5. Distortion by vertical tracking angle error.

Tracking Distortion

In the representations made in Figs. 1 and 2 it was assumed that playback was accomplished at right angles to the disc surface. This is usually not possible due to mechanical limitations, as shown in Fig. 5. As a result of the vertical tracking angle error, the playback stylus no longer moves identically to the recording on the disc. The maximum and minimum now occur at that point where the tangent at an angle δ to the disc surface touches the recorded waveform. There results a playback stylus motion as shown in Fig. 5: the rising flanks are condensed while the falling ones are expanded. This curve generates, aside from the fundamental, a second harmonic:

$$E_2 = E_1^2 \pi (1/V_g) \cos 2\omega t. \quad (3)$$

In an effort to remove the vertical tracking angle error problem at its source, disc recording standards worldwide have specified that recordings be made with a 15° vertical tracking angle and that playback cartridges also use the same angle. This is accomplished in the recording by tilting the cutterhead in the direction of groove travel. The amount of tilt is furthermore so chosen as to compensate lacquer spring back effects.

Groove Excursion Limit

The groove excursion limit at high frequencies is that excursion for which the stylus tip radius exactly equals the curvature radius of the modulation (Fig. 6). The equation for this level limit resulting from groove curvature is

$$V = V_g^2/R_s \cdot 2\pi f \quad (4)$$

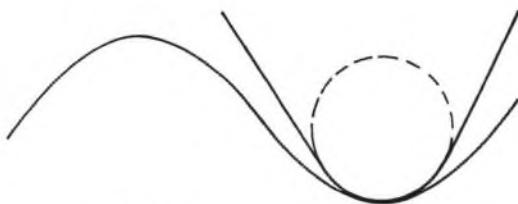


Fig. 6. Conventional recording.

where

V = velocity

V_g = groove velocity

R_s = stylus tip radius

f = frequency.

When playing back such a recording, there appears a second-harmonic distortion factor of 50%. Figure 7 shows very clearly that a predistorted recording under identical conditions produces a groove curvature which is significantly smaller. As a result, the groove excursion limit is reached at significantly higher levels.

II. CORRECTION POSSIBILITIES FOR GEOMETRICAL DISTORTION

Since it is expected that adherence to recording and playback standards will prevent distortion due to vertical tracking error, we must look to the compensation of tracing distortion for the next step in the improvement of the disc recording. In practice there are already two systems in use which serve this end. RCA Records utilizes in their Dynagroove system, a method first described by Fox and Woodward [1] in 1963. The dynamic recording correlator phase modulates the recording signal using a delay time as a function of the instantaneous value of the signal amplitude. Another system is used by Teldec (Telefunken-Decca) for the cutting of its Royal Sound stereo records. A tracing simulator described by Redlich and Klemp [2], [3] provides a means for compensating distortion by suitable addition of harmonics to the recording signal.

III. DESIGN OF A DEVICE FOR CORRECTION OF GEOMETRICAL DISTORTION

The Georg Neumann Company has developed a tracing simulator according to the method of Redlich and Klemp, which will be described in the following.

Block Schematic of Tracing Simulator

It was shown in Section I that the principal distortion occurs in the second harmonic. It is to be expected that up-to-date disc recording and reproducing systems adhere strictly to the 15° vertical tracking angle standard. The correction of the tracing distortion must be undertaken

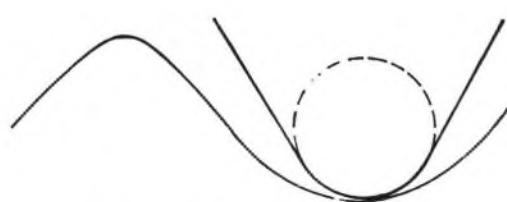


Fig. 7. Predistorted recording.

according to Eq. (2):

$$E_2 = E_1^2 \cdot R_s (1/V_g^2) \pi f \sin 2\omega t. \quad (2)$$

The schematic diagram of Fig. 8 is derived from this equation. The input signal E_i is fed to a squaring stage. Since the properties of such a squaring module are limited by its noise level and output capabilities, it is necessary to amplify the input signal to place it at the optimum operating point of the module. This is accomplished by the amplifier V and the level control k_1 . The signal $E_1 \sin \omega t$ arriving there is squared and then has the magnitude

$$-E_1^2 \sin^2 \omega t = -(E_1^2/2)(1 - \cos 2\omega t).$$

The dc voltage portion $E_1^2/2$ is eliminated through capacitative coupling. The factor E_1^2 from Eq. (1) has now been achieved, as has the required frequency doubling. The subsequent level control k_2 permits the signal to be brought into agreement with the factor R_s . While it is true that the stylus tip radius R_s is not within the control of the record producer, international standards have prescribed a radius of 0.6 mil, and the corrective signal must therefore be chosen for this value. Pickups deviating greatly from this standard cannot be expected to realize optimum distortion compensation.

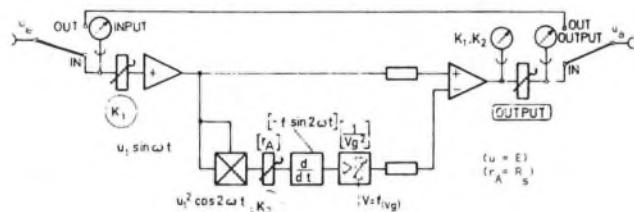


Fig. 8. Schematic block diagram for one channel.

In the next stage the signal is differentiated, and then has the value

$$(d/dt)[(E_1^2/2) \cdot R_s \pi \cos 2\omega t] = -E_1^2 \cdot R_s \cdot 2\pi f \sin 2\omega t.$$

Still remaining is the groove velocity control signal which must be taken from the carriage which transports the cutterhead, since it is only at that point that information regarding recording disc diameter and turntable r/min is available. The recording disc diameter is signaled by a series of microswitches which influence the gain of the amplification stage, which follows the differentiating stage. The turntable r/min can be inserted as a constant. We now have at the output of the amplifier the desired correctional signal

$$E_2 = -E_1^2 \cdot R_s (1/V_g^2) \pi f \sin 2\omega t \quad (1)$$

which is then fed to the inverted input of the combining amplifier V_2 . In the preceding analysis it has been assumed that the phase relationship between the power amplifier which follows the tracing simulator and the recording stylus is independent of the frequency applied, and is constant. Only under these conditions can the corrective signal produce the proper predistortion on the acetate record. Any phase deviations must be corrected electronically. While phase constancy may be obtained in the amplifier without difficulty, the same is not the case in the electromechanical transducer, since every sec-

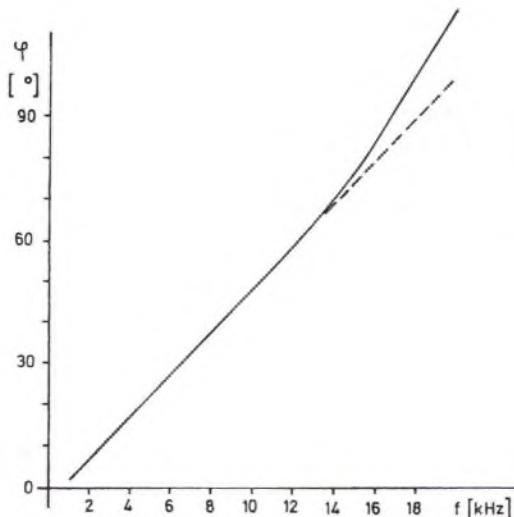


Fig. 9. Phase-frequency characteristic of cutterhead.

ondary resonance within the transmission range produces phase shift which cannot be compensated by level corrections. In the development of the stereo cutterhead, the Georg Neumann Company placed special emphasis on its phase linear properties. Figure 9 shows that within the passband this requirement was so well met that no phase correction is necessary for this cutterhead. Therefore the correction signal may be mixed in the proper proportion to the fundamental signal E_1 in a mixing amplifier V_2 . The output potentiometer permits adjustment to zero insertion loss.

Structurally the tracing simulator is divided into individual function groups augmented by a test and calibration module. The use of operational amplifiers and mathematical function modules provides for structural simplicity.

Fundamental Frequency Module Construction

A single plug-in card mounts all of the components necessary for the processing of the fundamental in both channels of the tracing simulator (Fig. 10). The input signal E_i is first fed to a potentiometer K_1 , K_1 being the fundamental. It is to be adjusted such that for a given condition (for example, peak recorded velocity in each channel of 8 cm/s at 3150 Hz) the ensuing quadratic function generator operates in its optimum working range.

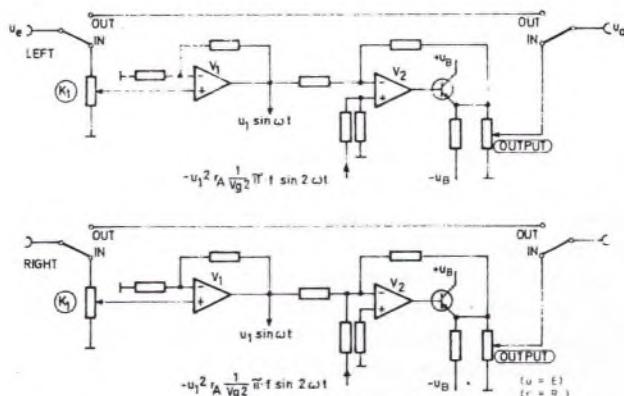


Fig. 10. Schematic block diagram for circuit for fundamental.

The level diagram of the transfer system in question requires subsequent amplification of 25 dB.

The fundamental and the corrective signal are subsequently combined in a mixing amplifier V_2 . This mixing amplifier has a loss of 20 dB for the fundamental and of 10 dB for the harmonic. The output potentiometer is used to adjust for zero insertion loss. The corrective signal is mixed into the left channel in antiphase and into the right channel in-phase. The reason for this is to be found in the fundamental structure of a stereo cutterhead. According to international convention a stereo cutter should produce lateral motion for in-phase signals fed to it. This requires that one channel perform a motion into the disc while the other retracts the stylus from the disc. Because of this motional process the corrective signals (2) must be mixed into the fundamental in opposing phase relationship (Fig. 11).

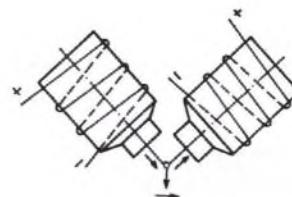


Fig. 11. Stereo cutterhead function.

X² Function Module Construction

There are various possible methods for forming the square of the input signal. The stringent requirements, however, reduced these considerably. The operating range must be extensive since the dynamic range is expanded through squaring. (30 dB dynamic range in the fundamental results in 60 dB after squaring). This requires a quadratic function generator of high output capability and low self-noise level. Throughout the entire frequency range the quadratic function must be precise, the output capability must be assured, and the phase relationship between fundamental and harmonic must agree.

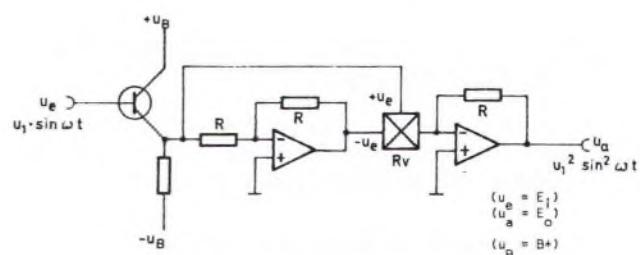


Fig. 12. Schematic block diagram of squaring circuit.

The quadratic range of rectifiers is too small. Hall generators require excessive circuitry for feeding the control winding and may produce difficulties in the maintenance of the phase requirements. Multiplier circuitry from the analog computer technology utilizing the current distribution principle, while they have sufficient output capability, display a self-noise level two orders of magnitude too poor. The most useful solution was found in the parabolic multiplier which has both sufficient output capability and low noise level. Circuitry according to Fig. 12 was used in the tracing simulator. The squaring module works as an input resistor in front of the inverted

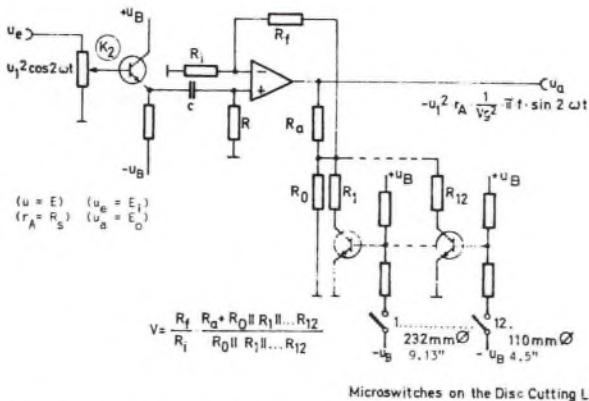


Fig. 13. Differentiating circuit and remote-controlled amplifier.

input of an operational amplifier. Its source impedance is inversely proportional to the applied signal voltage $R_v \sim R_o/E_i$. As a result there is obtained at the output of the amplifier the squared signal

$$\begin{aligned} E_o &= -E_i \cdot V \\ &= -E_i \cdot R_2/R_v \\ &= -E_i \cdot R_2/R_o \cdot E_i \\ &= \text{constant } E_1^2 \sin^2 \omega t \\ &= \text{constant } E_1^2 (0.5 - 0.5 \cos 2\omega t). \end{aligned}$$

When the signal is missing or very small, the value of the source impedance is very large, the amplification of the operational amplifier is small, and as a result the output noise level is small too.

Further Processing of Second Harmonic

Figure 13 shows further processing of the squared fundamental signal. The potentiometer K_2 adjusts for the stylus tip radius of 15 μm (0.6 mil) and also allows for compensation of component tolerances.

Following an impedance converter, the signal is differentiated by means of the capacitor C and resistor R with a boundary frequency of 160 kHz. The deviation from the ideal 90° phase angle at the upper limit of the passband (20 kHz) is only 7°. The gain of the ensuing operational amplifier is variable, allowing the corrective signal to be made a function of linear groove velocity $1/V_g^2$. Its feedback is determined by a series of resistors R_1, \dots, R_{12} which are added to the circuit by means of transistor switches. The microswitches are mounted on

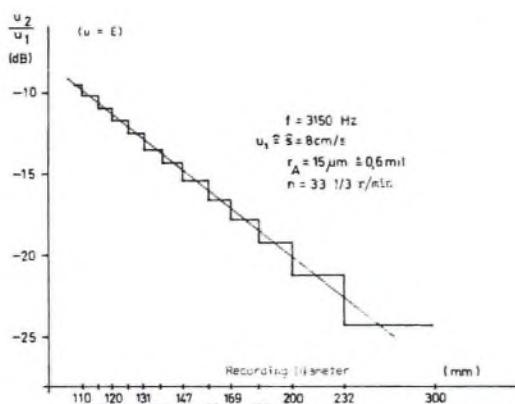


Fig. 14. Relation between second harmonic and recording diameter.

a disc-cutting lathe and are operated by the carriage moving the cutterhead across the disc.

The dependence of the corrective signal on the groove velocity $1/V_g^2$ is therefore approximated by a step function curve. The number of steps used (12 steps) and the ever closer spacing of the steps as the diameter is reduced provide for a sufficiently accurate reproduction of the desired curve (Fig. 14). Compensation is effective beginning with the outermost groove on the disc. We therefore have at the output of the amplifier the desired corrective signal for insertion into the fundamental card:

$$E_2 = -E_1^2 \cdot R_s (1/V_g^2) \pi f \sin 2\omega t.$$

Calibration Module

A further plug-in card mounts the elements necessary to facilitate the calibration of the tracing simulator. The following requirements must be met by such a calibrating module:

- 1) The insertion loss of 0 dB must be checkable (potentiometer input and output).
- 2) The adjustment of the proper level for the quadratic unit must be measurable (potentiometer K_1).
- 3) The proper amount of second-harmonic injection must be measurable (potentiometer K_2).

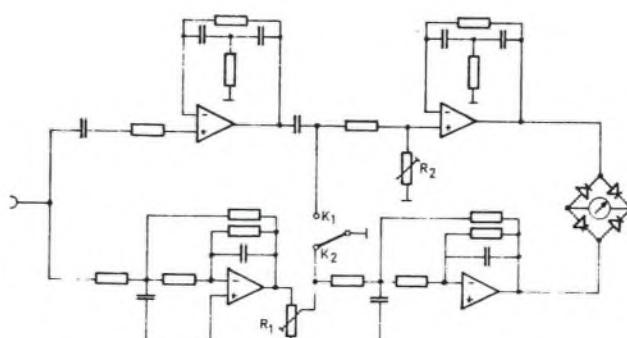


Fig. 15. Test circuit.

To permit calibration, a nominal alignment situation must be chosen: frequency 3150 Hz, peak recorded velocity per channel 8 cm/s, groove velocity corresponding to a diameter of 108 mm. The choice of this parameter produces measurement voltage signals of a reasonable order of magnitude. The recorded level lies 3.4 dB below the maximum recordable velocity, and the second-harmonic portion of the signal is 9.5 dB below the fundamental. To measure the harmonic, a Butterworth band-filter of the fourth order was chosen, which for steep enough flanks has a sufficiently flat peak (Fig. 15). This assures that a slight deviation of the test frequency does not falsify the test results while the three times stronger fundamental is suppressed.

For measurement of the fundamental a Butterworth low-pass filter of the fourth order is used. The test meter is connected between the outputs of the two test amplifiers. Depending on the test frequency, a simple switch connects the unused test circuit, and with it the test meter, to ground. By means of the potentiometers R_1 and R_2 the sensitivity of the two test-amplifier branches may be chosen so as to produce a predetermined pointer deflection for proper calibration.

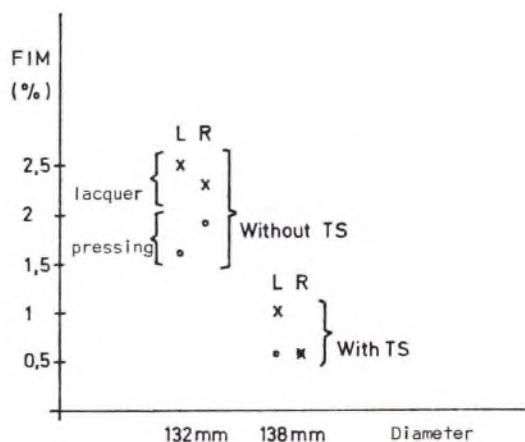


Fig. 16. Measured distortion on lacquer and pressing.

PRACTICAL OPERATION WITH TRACING SIMULATOR

The tracing simulator provides another step in the improvement of disc recording. As a result of this improvement we now can determine other defects in the disc transfer chain which have heretofore not been noticed. This includes the entire chain from the tape to the disc playback.

Choosing a Test Method

The addition of harmonics in the tracing simulator is performed just as the theoretical calculation requires. The question now is to devise a test with which may be proven whether the expected distortion compensation actually takes place. The obvious possibility of performing a selective measurement of the second harmonic was shown to be unsuitable. The playback process results in geometric distortion resulting from frequency and phase modulation, but this also results in harmonic distortion products. Nonlinear behavior of the transducer and tracking distortion as a result of improper guidance of the stylus in the groove produce harmonic distortion which has no relationship with the geometrically caused distortion. These distortion products are not compensated by the tracing simulator, remain as residual distortion, and give the impression of insufficient compensation. Practical experimentation has shown that measurement of the harmonic distortion products do not lead to a determination of the quality of the recording, but rather of the insufficiency of the transducer. Measurement of the frequency

modulation, on the other hand, encompasses the geometric tracing distortion principally. For this test a two-frequency recording is suitable, whose frequency intermodulation (FIM) may be measured by simply using an ordinary wow and flutter meter. In order to utilize standard test methods, a frequency pair of 1:10 ratio (carrier 315 Hz, modulating frequency 3150 Hz) is chosen with recorded velocities of a ratio of 4:1.

When deciding on a distortion measurement method it must be borne in mind that the test results should correspond to the psychoacoustic impression. Practical experiments have shown that the annoyance factor of distortion resulting from record playback using various pickups correlates well with their FIM value.¹ Using a test method based on evaluation of the harmonic distortion or amplitude modulation value produces results which do not correlate in such a way. Therefore the measurement of the FIM of a frequency pair with and without the use of the tracing simulator provides a good method of checking the compensation.

Demands of Cutterhead, Stylus, and Acetate Blank

In Section III it was shown that the cutterhead must have a constant phase response over the entire frequency range. In the investigation of distortion caused by the vertical tracking angle (Section I) it was determined that such distortion can be avoided altogether by making sure that the vertical tracking angle for the cutterhead and playback cartridge agree. The vertical tracking angle of the recording is dependent on the geometry of the cutterhead and the spring-back properties of the lacquer used in the acetate blank. While the geometry of the cutterhead is a known factor, the spring-back property of acetates is caused by numerous factors. It is primarily, of course, dependent on the materials used in manufacture, and therefore blanks of different manufacturers may produce differing spring-back angles. Another factor is the stylus temperature, and the shape and evenness of the stylus facettes. The subsequent processes through which the master passes may also contribute errors. It should therefore be the aim of the industry to determine and hold constant these parameters, since deviations may adversely influence the degree of compensation attainable with the tracing simulator.

The analysis of test values with and without the tracing

¹ Information based on investigations at Teldec Laboratories [4].

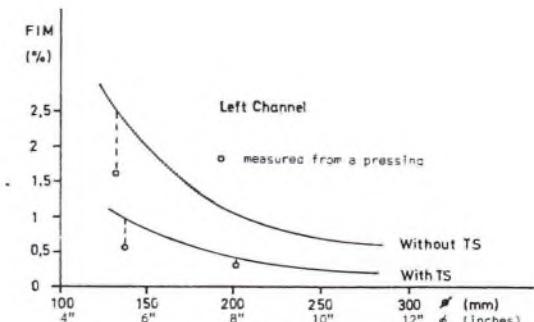


Fig. 17. Distortion reduction.

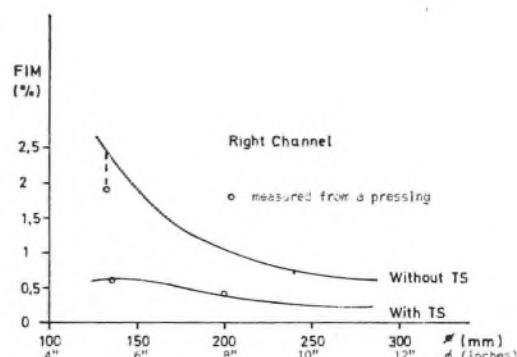


Fig. 18. Distortion reduction.

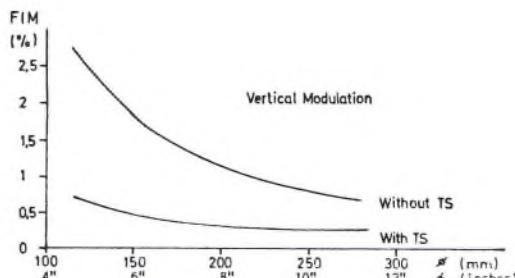


Fig. 19. Distortion reduction.

simulator can be used to determine the wear of styli, since initially measured FIM values with a new stylus will deteriorate through stylus wear. Poor FIM values with a new stylus may indicate an unsuitable stylus, or they may indicate that the stylus is not properly seated in the cutterhead's stylus chuck. A loose fit may cause the stylus to perform motion in the direction of groove travel, causing frequency modulation of the recorded signal. This in turn yields poor FIM test results (tracing distortion), but cannot be compensated with the tracing simulator.

Demands of Playback Cartridge

Should the tracing distortion of a recording and its compensation be proven through playback tests, it must be ascertained that test results will not be falsified by other distortion sources. This places very stringent demands on the pickup if the measurement is to be made from acetate playback.

- 1) The vertical tracking angle must adhere to the standard of $15^\circ \pm 3^\circ$. Furthermore it must be independent of frequency and attainable at the optimum stylus force.
- 2) The stylus tip radius must conform to the standard of $15 \mu\text{m}$ (0.6 mil), since optimum compensation can otherwise not be attained.
- 3) The pickup must make good contact with the

groove walls. Satisfying this requirement seems to present great difficulties. Tests have shown that minimum distortion can only be attained with a stylus force at or even beyond the upper limit suggested by the manufacturer. The skating force must be compensated to assure equal contact with both groove walls. Even when this requirement is satisfied, differences are noticed if various styli are used with the same cartridge. It seems that resonance effects between stylus and lacquer seriously influence distortion tests. Tests on acetates show that even when the skating force is compensated, there is a noticeable difference in the residual distortion of both groove flanks, while tests on pressings no longer show such differences (Fig. 16). It is therefore not necessarily so, that the values obtained from acetate playback can be duplicated when playing the pressing from the same acetate. The pressing may be better.

The compensated to uncompensated ratio (compensation factor) does remain constant when measurements and equipment are carefully maintained, so that acetate playback measurements will permit determination of geometric distortion compensation through the tracing simulator. Figures 17-19 show the results of such tests. The X show comparative values obtained from a pressing.

REFERENCES

1. E. C. Fox and J. G. Woodward, "Tracing Distortion—Its Cause and Correction in Stereodisc Recording Systems," *J. Audio Eng. Soc.* **11**, 294 (1963).
2. H. Redlich and H. J. Klemp, "Tracing Simulator—Ein Verfahren zur Schallplattenaufzeichnung für verzerrungssarme Wiedergabe," *Funktechnik*, 161 (1965); also *Elektron. Rundschau* **19**, 1, 15 (1965).
3. H. Redlich and H. J. Klemp, "A New Method of Disc Recording for Reproduction with Reduced Distortion: The Tracing Simulator," *J. Audio Eng. Soc.* **13**, 111 (1965).
4. H. Redlich and H. J. Klemp, Teldec Laboratories, Berlin, Germany, private communication.

THE AUTHOR



Dieter Braschoss was born in Spremberg, Germany, in 1932. He received the Dipl.-Ing. degree from the Technical University, Berlin, in 1956.

After having worked in the Component Section of

Siemens AG, Berlin, for 8 years, Mr. Braschoss joined Georg Neumann GmbH, Berlin, in 1964 as a Development Engineer. At present he is heading the development staff for disc cutting techniques.

Groove Echo in Lacquer Masters*

DANIEL W. GRAVEREAUX AND BENJAMIN B. BAUER

CBS Laboratories, Stamford, Conn. 06905

The audio signal which is transferred from modulated grooves to adjacent unmodulated grooves in a disc record is called the "echo." The measured characteristics of the echo signal are described as a function of modulation level, modulation frequency, groove width, and cutting pitch (grooves per inch). It is shown that echo during the mastering process is produced by the modulation of the left (inner) groove walls and is transferred to the succeeding inner grooves.

INTRODUCTION: Disc and tape records often contain audio information at reduced levels preceding or following the actual audio signal. Since this information is similar to the main signal, it is called an "echo signal." When audible, the echo signal is a cause for quality control rejection of manufactured records. This paper is a report on a study of disc groove echo.

There are many causes for the occurrence of disc groove echo, one of them being the echo which sometimes is present in the master tape, caused by the transferral of magnetic flux from one layer of tape to another in the case of reel-wound magnetic tapes. This latter phenomenon is termed "print-through." The disc record sometimes is erroneously blamed for the echo in the master tape and, fortuitously, one turn of a 10 in reel of tape at 15 in/s corresponds approximately to 1 turn of the disc. In this paper we are not concerned with this latter problem but only with the echo caused by the actual process of cutting the lacquer disc. We examine the character of this echo and determine the effects of various parameters upon the level of the echo signal.

There is some evidence based upon experience of operating personnel that echo in the final pressed disc may be influenced by processing operations following the cutting of the lacquer master, such as the temperature of the plating bath, etc. These variables were not a part of this investigation which was directed at the question of echo formation during the lacquer cutting process.

In our investigation, a series of recordings was made on lacquer masters, and echo was measured as a function of (1) the groove geometries, (2) the modulating frequency, and (3) the playback stylus tip size.

MEASUREMENT TECHNIQUE AND EQUIPMENT

The echo was measured by modulating the lacquer groove for the duration of one turntable revolution and measuring the recorded modulation level as well as the echo level in the preceding and following turns of the groove.

Fig. 1 describes the recording and playback equipment. The oscillator, with an output attenuator calibrated in dB referred to disc velocity, is connected to the contacts of the timer. Its output is then delivered to either the left or right channels of the disc-cutter amplifier. The disc-cutter equalization is described by a

* Presented April 27, 1971, at the 40th Convention of the Audio Engineering Society, Los Angeles.

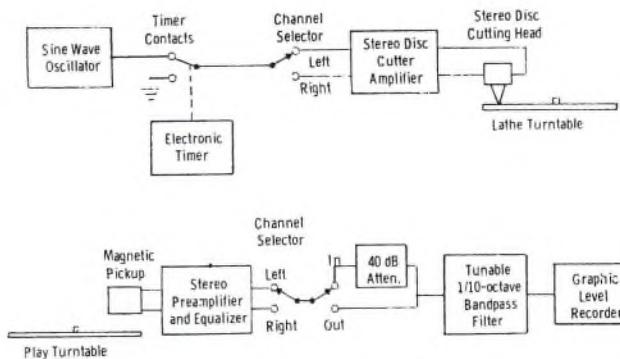


Fig. 1. Equipment setup for performing echo measurements.

single RC network with a time constant of $312 \mu\text{s}$, providing asymptotically a constant amplitude response below 500 Hz and a constant velocity response above 500 Hz. A Westrex 3D stereo cutter head is used. A $160\times$ microscope with a calibrated reticule mounted on the lathe permits accurate measurement of groove width and spacing between the grooves.

The playback system utilizes a good quality commercial turntable and a medium-priced magnetic pickup with various-sized stylus-tip assemblies, although a later experiment showed that echo is independent of the playback stylus radius. The pickup output may be amplified with a preamplifier having either a flat or an equalized transmission characteristic. The equalization consists of a single RC network with a time constant of $312 \mu\text{s}$ such that when driven by a velocity-responsive pickup, it delivers constant voltage at all frequencies when playing back a disc mastered with the above equalized cutter. The equalized playback characteristic was found to be convenient to use in investigating the frequency dependence of the echo. Either the left or right channel is fed into the measuring system which consists of a switchable 40-dB attenuator, a tunable 1/10-octave wide bandpass filter, and a graphic level recorder.

In order to obtain reliable and rapid measurements, an automatic timer was used to cut the single modulated revolution. The signal and the echo levels during playback were traced on a strip chart utilizing a graphic level recorder. The main signal was attenuated by 40 dB during the modulated groove portion, and then returned to the unattenuated position in the echo portions of the disc. This method facilitates reading the echo data with reference to the fundamental signal level and prevents overloading the bandpass filter. The tunable filter passes the sinusoidal signal while rejecting hum, rumble, and random noise, so that groove echo values of 80 dB below 3.54-cm/s velocity can be measured. The 1/10-octave filter was always tuned to the frequency of the signal (or the echo signal) being measured.

MEASUREMENTS PERFORMED AND DISCUSSION

Nature of Echo

In order to examine the general characteristics of the groove echo, a band was cut on a lacquer disc at diameters of 11 to 10 in, at $33 \frac{1}{3}$ r/min, with the following program. A 1-kHz sine wave, 5 dB below 3.54-cm/s rms velocity, was modulated on the left channel for one revolution of the disc every four revolutions.

This was repeated for the right channel. Cuts were made with the pitch set at 450 lines per inch, at groove widths of 1.6 and 2.2 mil. It should be noted that this latter width corresponds to a "no-land" condition.

Qualitative examination of the echo levels upon lacquer playback showed the following characteristics.

1) Referring to Fig. 2 which depicts the experiment, it is seen that echo occurs mainly as postecho; that is, the signal is transferred predominantly from the signal-unmodulated groove to the adjacent inner groove. Echo in the already cut grooves is at quite a low level (60–70 dB below the signal level).

2) The wider groove resulted in considerably increased values of postecho.

3) The postecho is mainly created by modulation in the left channel of the preceding groove. Modulation in the right channel produces postecho readings of at least 60 dB below the main signal.

4) The postecho signal appears equally (within ± 1 dB) in both the left and right channels of the following groove.

Thus the mechanism of the echo formation appears to be as follows. The process of cutting the groove affects the stress distribution in the lacquer, and this effect evidently is greater in the relatively solid mass of uncut lacquer following the groove than in the more compliant portion between the groove being cut and its preceding turn. This stress distribution is relieved when the succeeding turn is cut, resulting in the formation of echo.

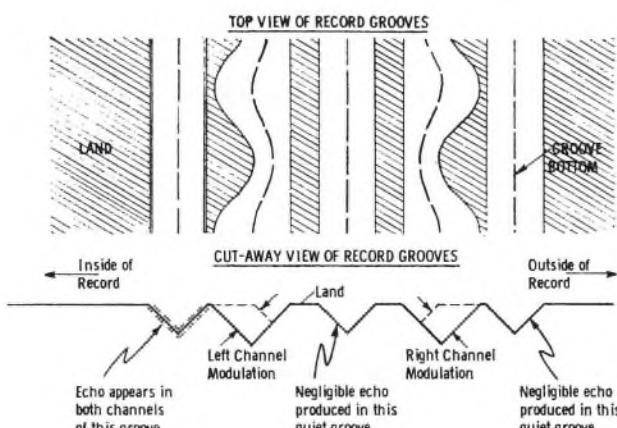


Fig. 2. Nature of echo versus record geometry.

Echo as a Function of Fundamental Level

Our next set of measurements determined the relationship of the recorded amplitude upon the level of the echo signal. A band from $8 \frac{1}{2}$ to 8 in diameter was recorded with a pitch of 450 grooves per inch and with a groove width of 2 mil. A 1-kHz signal applied to the left channel was switched on for one revolution out of four, and repeated with the input signal reduced by 3 dB each time so that the signal varied from 0 to -15 dB referred to 3.54 cm/s rms.

Playback of the lacquer disc showed that the postecho caused by left channel modulation in the previous groove remained 36 dB below the main signal level at all times. This suggests that the echo is a linear transference, at least for this wavelength and groove geometry.

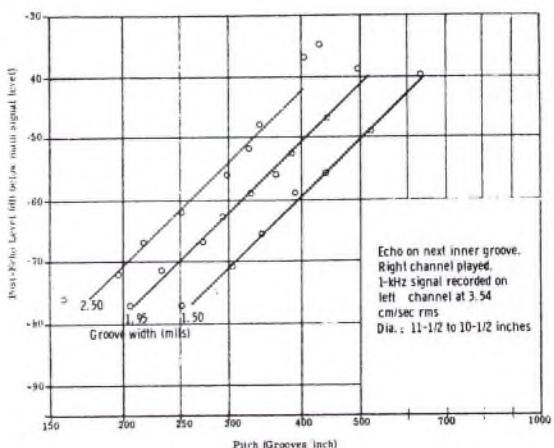


Fig. 3. Postecho versus recording pitch (right channel played).

Echo as a Function of Recording Pitch

A set of measurements was performed to obtain the relationship of distance between the grooves to the level of postecho. This distance was defined in terms of pitch, or grooves per radial inch, in order to present the results in terms of recording vernacular. A band was cut between 11 1/2 and 10 1/2 in in diameter while recording the 1-kHz 3.54-cm/s rms sine wave on the left channel. One revolution out of four was modulated in repetition. Three different groove widths were used: 1.50, 1.95, and 2.50 mil. In each case the pitch was varied in discrete steps covering a range of about 3 to 1.

Postecho levels were obtained by playing back both the left and right channels. The pitch was measured by counting the number of grooves in a 22-mil wide microscope field. Fig. 3 and 4 are graphs portraying the result: echo level in dB versus log pitch. The data points fall well along a straight line (± 1.5 dB) for all levels, except where the echo-bearing groove cuts into the modulated groove. In general a postecho level of about -40 dB occurs for just overcut grooves regardless of groove width. Decreasing the pitch by a factor of two will reduce postecho by approximately 28 dB or 25 times.

Echo as a Function of Groove Width

Fig. 3 and 4 showed that postecho is also a function of groove width. In order to further establish this relation-

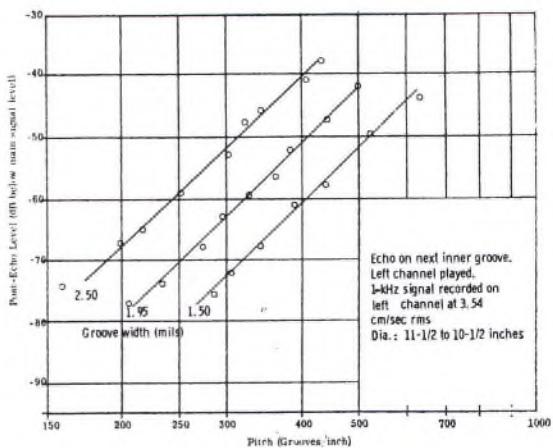


Fig. 4. Postecho versus recording pitch (left channel played).

ship, a band was recorded at 11 1/2 to 10 1/2 in in diameter with the groove width as the parameter. The 1-kHz sine wave, -3 dB referred to 3.54 cm/s rms, was switched on and off for one revolution out of four. It was applied to the left channel only. The groove width was adjusted for 3.4 mil and then decreased in uncalibrated, but discrete, steps every four revolutions of the lathe turntable until a width of 1.35 mil was reached.

Both the modulated groove level and the left and right channel echo levels were read on playback. The microscope was used to measure the groove width. The plot of the data points is given in Fig. 5. Both the left and right echo levels are nearly the same. It is seen that a reduction of echo level of about 24 dB or 16 times can be obtained by reducing the groove width by a factor of 2. The two points appearing above the straight line are the right channel echo for overcut grooves.

Echo as a Function of Signal Frequency

All previous data were obtained for a 1-kHz sine wave modulation. To ascertain the effect of frequency, or wavelength, the postecho was measured for discrete frequencies from 100 Hz to 15 kHz at both the inside and outside diameters of a disc. Two bands were cut on a lacquer disc, one from 11 in to 10 1/2 in in diameter and another from 5 1/2 to 5 in in diameter with a pitch of 333 lines per inch and a groove width of 2.4 mil. The groove was modulated one rotation out of four with the

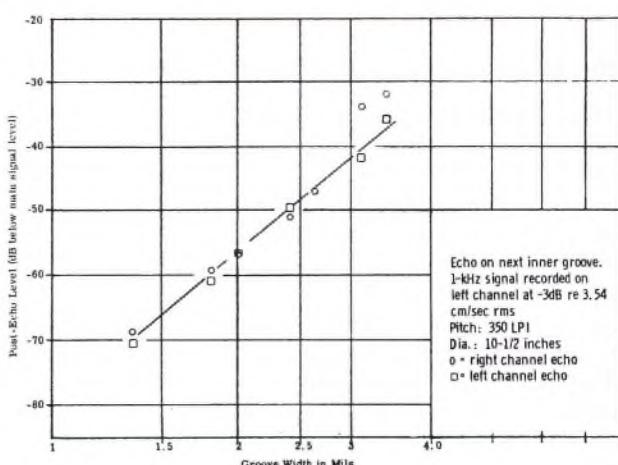


Fig. 5. Postecho versus groove width.

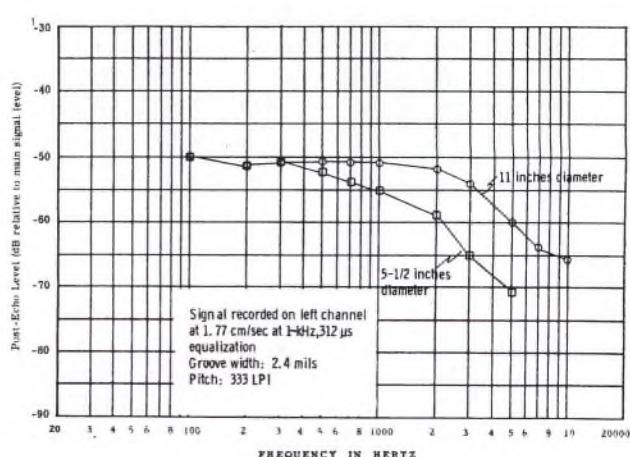


Fig. 6. Postecho versus frequency.

signal oscillator frequency raised every four revolutions to cover the frequency range. The signal was applied to the left channel only at a level of 1.77 cm/s at 1 kHz.

The lacquer disc was played back through the equalized preamplifier, and the magnitude of postecho, relative to the main signal, was measured. Fig. 6 shows the levels of echo on the right channel of the next inner groove. We observe that the postecho at both the inside and outside diameters of the disc are the same at low frequencies. However, the high-frequency echo is a function of diameter, implying wavelength dependence. At the outside diameter, the echo diminishes with frequency at a rate of 12 dB per octave, with an intercept at 2 kHz. At the inside diameters the echo decreases at about the same rate with an intercept of 1 kHz.

Echo as a Function of Playback Tip Size

A number of the above measurements were performed utilizing three different playback tip assemblies, a 0.7-mil-radius spherical, a 0.5-mil-radius spherical, and a 0.3- by 0.7-mil elliptical tip. The postecho level remained essentially constant regardless of tip dimensions.

Echo Equation

From the aforementioned experiments the following empirical equation is derived:

$$\frac{E}{S} = K_1 \left[W_g^4 P^{4.64} \left\{ \frac{1}{1 + K_2/\lambda^2} \right\} \right]$$

where

E = echo level

S = signal level

K_1, K_2 = constants

W_g = groove width

P = pitch

λ = mechanical wavelength.

The constants have been determined from the data for the case of groove width given in inches $\times 10^{-3}$

(mils), pitch given in grooves per inch, and wavelength given in centimeters:

$$K_1 = 1.85 \times 10^{-16}$$

$$K_2 = 7.$$

Converting from wavelength to turntable rate, disc diameter, and frequency, and performing the operation $20 \log E/S$, we obtain the following postecho level in dB referred to the main signal level:

$$E_{dB} = 80 \log_{10} W_g + 93 \log_{10} P - 20 \log_{10} (1 + 1.25 \times 10^{-2} f^2 / D^2 N^2) - 315$$

where

E_{dB} = post echo level in dB relative to fundamental signal

P = pitch in grooves per inch

f = signal frequency in hertz

D = disc diameter in inches

N = turntable rate in r/min

W_g = groove width in inches $\times 10^{-3}$.

CONCLUSIONS

Groove echo, occurring during cutting a cellulose-acetate lacquer blank, occurs mainly as postecho, with the audio information being transferred from the inner groove wall (left channel modulation) to the next adjacent groove where it occurs in equal amounts at both groove walls. Echo is generally negligible for wide-pitch narrow-groove recordings; however, it increases rapidly in close-pitch high-density records. Below 1-2 kHz, echo is independent of frequency. At higher frequencies, the echo signal diminishes inversely with the square of the frequency and is further diminished toward the inside diameter of the disc. Preecho, occurring during "lead-in" or in "crossover" bands, usually is not caused by the lacquer mastering process.

Note: Mr. Gravereaux's biography appeared in the September issue and Mr. Bauer's in the June issue.

LETTERS TO THE EDITOR

COMMENTS ON "GROOVE ECHO IN LACQUER MASTERS"*

D. A. BARLOW

Bingley, Yorks., England

The authors found that preecho, i.e., echo from the following groove, was very low; postecho, i.e., echo from the previous groove, could be quite high. They suggest that this is a function of the changing of the internal stress system when the following groove is cut. The present writer suggests that the reason is simply the normal springing of the groove wall under the cutting stresses. This may be illustrated as follows.

In Fig. 1, as cutting proceeds, the unmodulated groove *A* is cut. On the following revolution, the modulated groove *B* is cut. Where the modulation approaches the wall of the previous groove at *C*, the thin wall will spring or elastically deform under the lateral force, as shown by the dotted line. After the cutter has passed, the left wall of groove *A* will spring back to its original position, or nearly so. The right wall of groove *B* will spring back to the dotted line position as indicated.

Postecho is illustrated in Fig. 2. As cutting proceeds, the modulated groove *A* is cut. After another revolution, the unmodulated groove *B* is cut. At *C* the wall between the grooves will deflect under load from the cutter, as indicated. After the cutter has passed, this will spring back, pushing in the wall of groove *B*, as shown.

In each case, the already cut groove is not affected, i.e., there is no preecho, but the groove being cut is af-

fected, i.e., there is considerable postecho. Any bulk plastic deformation of the material between grooves will give preecho and reduce postecho. The residual stress system left in the lacquer adjacent to a cut groove will interact with that of a closely spaced following groove. This may modify the echo but will be a second-order effect compared with spring of the groove walls under load.

The existence of postecho on the *L* wall, of similar magnitude to that on the *R* wall, may also be explained by the springing of the groove wall. The cutter will take the line of least resistance, in spite of the input signal. If there is less material at one side, it will drift to that side, i.e., the *R* side for postecho. The resulting profile of the *L* wall being cut will be as shown in Fig. 2, i.e., there will be postecho on the *L* wall also. This can be checked by comparing the phase of the postecho on each wall.

The influence of wavelength at high frequencies is

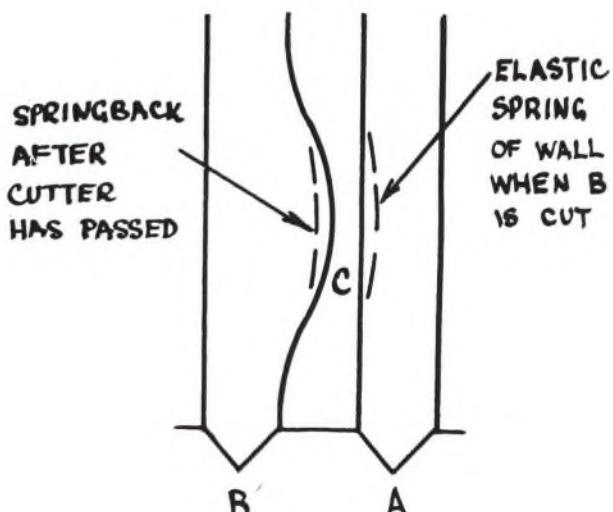


Fig. 1. Preecho.

* D. W. Gravereaux and B. B. Bauer, *J. Audio Eng. Soc.*, vol. 19, pp. 847-850 (Nov. 1971).

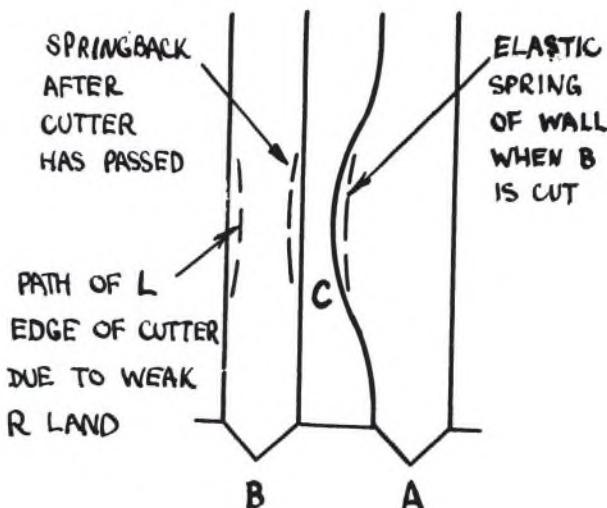


Fig. 2. Postecho.

shown in Fig. 3 for a given frequency, amplitude, and groove spacing. With the longer wavelength there is a greater area of weak wall, so that postecho will be greater at the outside than the inner grooves of the record. At low frequencies, when the wavelength is very long, the two grooves are almost parallel. The spring of the wall and drift of the cutter will be greater than for short wavelengths, and will be independent of wavelength if this is sufficiently large (Fig. 4).

It follows also that for a given shape of groove, regardless of size, the postecho will be similar for a given frequency, amplitude, and wavelength at the point of just overcutting (as in Figs. 3 and 4 of Gravereaux and Bauer*). Likewise, for a given frequency, amplitude, wavelength, and groove spacing, wider grooves will give greater postecho simply because the land between grooves

is less. Elastic deformation of the material is by definition linear, although the effect of geometry may not be so; nevertheless it is not surprising if the overall effect is linear.

It would be expected that stylus radius would not affect echo level. The deformations due to the various styli will be different, but the relative differences, i.e., the signal levels, will not. If the trace radii were very small, the various styli might give different results, but this obviously does not apply here, as levels are very low.

In processing, if the plating is done hot, some relief of residual stresses in the lacquer might take place, and this would affect the echo to a small extent.

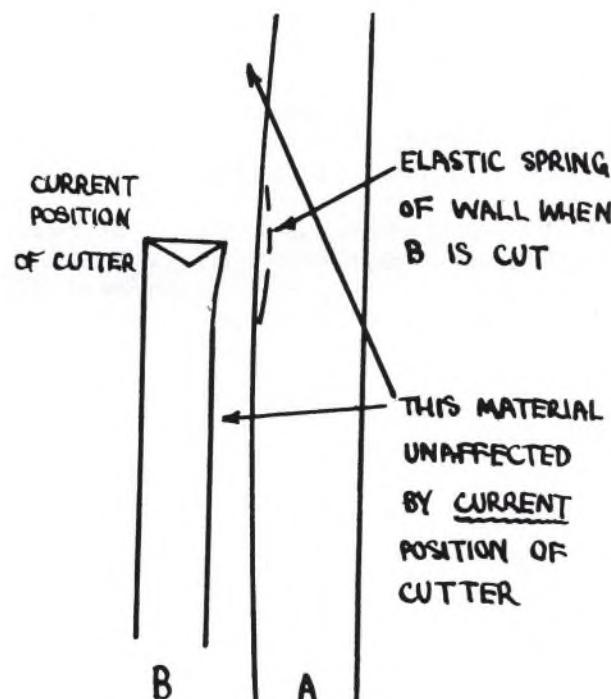


Fig. 4. Low frequencies.

Reply by Gravereaux and Bauer

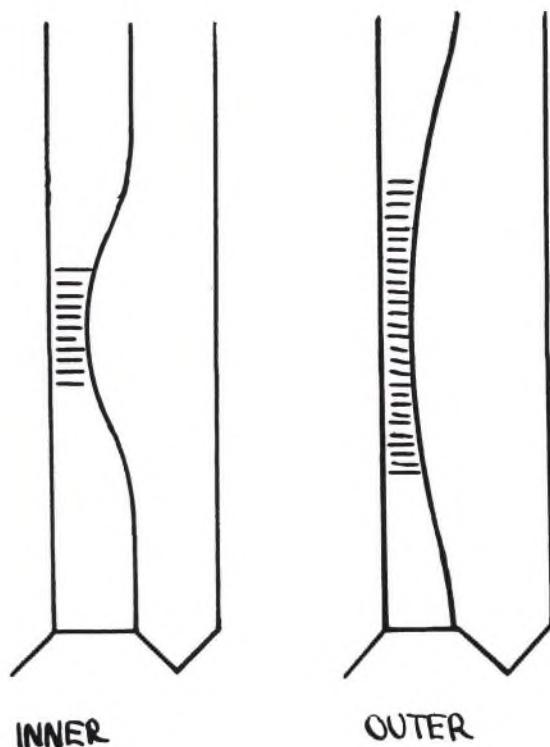


Fig. 3. High frequencies.

While we cannot claim academic rigor respecting our theory of groove echo formation, we believe that our explanation is closer to the truth than Mr. Barlow's. He suggests that echo is produced by strain induced in a first groove by the action of the cutter which is relieved when the following groove is cut, resulting in an in-phase modulation at the groove wall nearest to the first-cut groove and, by reaction of groove upon the cutter, in an out-of-phase modulation at the wall remote from the first-cut groove. Our theory from lacquer stress considerations suggests that the echo should be out of phase with the primary modulation at both the near and the remote walls. In other words, the echo modulation should be lateral, rather than vertical. Proving that echo is out of phase with the modulation of the first groove is a relatively difficult task since it requires the simultaneous measurement of both signals. It is easy to show, however, that echo modulation in the following groove is a lateral

modulation. A simple measurement proves that vertical modulation in an echo-bearing groove is nil, thus disputing one of the hypotheses of Mr. Barlow.

Consider our theory of groove echo formation in a little more detail. Everyone is familiar with stresses set up in a lacquered surface as it dries. An extreme example of this is crazed or checkered lacquered furniture, which indicates that the lacquer has shrunk as a result of drying until stresses have exceeded the tensile strength of the material. The surface of the disc lacquer is under such stress, which, however, is considerably less than the tensile strength of the lacquer, thus the surface remains in equilibrium until the first groove is cut. Let this first groove be modulated with the left (or inner groove-wall) channel signal. The lacquer surface becomes displaced laterally, causing a slight widening of the groove, with the strain lines on the disc surface remaining parallel with the groove modulation for a considerable radial distance. As several succeeding blank grooves are cut the strain is relieved and the whole mass of material is allowed to spring back with the result described in our paper.

A key piece of evidence is missing. Is the echo modulation in, or out of phase with the echo-causing signal? Perhaps someone will perform this measurement and shed more light on the phenomenon of echo formation.

Development of Skew-Sampling Compensator for Tracing Error*

SHIGETAKA WASHIZAWA, TOMOFUMI NAKATANI, AND
TAKEO SHIGA

Nippon Columbia Company, Ltd., Kawasaki 210, Japan

Tracing error is prevented from appearing in disc-record playback if the skew-sampling compensator proposed by Cooper is used in recording. Such a machine has been developed. Its specifications and principles of operation in relation to the skew-coordinate theory of tracing error are described. Examples of experimental data showing achievable distortion reductions are given.

INTRODUCTION: In the playing of a disc record, several distortions occur as the result of a stylus tip that does not correctly trace a signal in the groove.

The skew-sampling method is one proposed by Cooper [2] for removing these errors. In this method, delay modulation is introduced in the original signal upon recording to prevent the appearance of tracing error in the playback waveform. The authors have developed a compensator in accordance with this theory. Experimental data are shown in this paper indicating that excellent results in error compensation are obtained by this method.

TRACING DISTORTION

The distortions from tracing and tracking errors have been shown to arise from a delay modulation of the signal, for which the instantaneous delay variation is directly controlled by the modulated signal itself because of the geometrical relation between a stylus tip and a groove wall. By tracing distortion is meant the fact that

the sound groove of a disc record, cut with a V-shaped cutting stylus tip right and left as well as up and down in correspondence with the recording signal, has such a character that in order to produce the signal by means of a pickup stylus having a tip of semispherical form, the locus of the vibration traced by the pickup stylus is different from the locus cut by the cutting stylus.

Fig. 1 shows the representation of a spherical stylus tracing a sinusoidal modulation in the plane of modulation of one of the 45/45 channels. As shown in this figure, the locus of the center of the stylus is different in shape from the sound-groove wall, such that when the radius of curvature at the stylus tip approaches that of the waveform of the sound groove or the slope of the waveform becomes steeper, the tracing error Δx between the center of stylus and the stylus-groove contact point increases, and the center of stylus is caused to trace a distorted locus different from the modulation of the groove wall.

The tracing distortion may be removed. If an original signal is delay modulated to exactly cancel the delay variations Δt caused by these errors Δx in the velocity

* Presented April 27, 1971, at the 40th Convention of the Audio Engineering Society, Los Angeles.

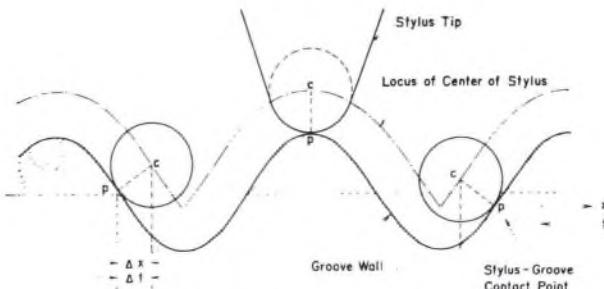


Fig. 1. Spherical stylus tracing a sinusoidally modulated groove wall.

signal, and a groove wall is formed as shown in Fig. 2, such that the locus of the center stylus traces the same locus as the groove wall in Fig. 1, the original signal.

When the groove velocity is V_g , the velocity of the modulated signal is V , and the stylus radius in playback is r , the instantaneous delay Δt corresponding to a tracking error Δx is given by

$$\Delta t = rV/V_g^2. \quad (1)$$

Skew-sampling method is the method in which a playback error is estimated in an original signal in accordance with geometrical relationship of the stylus profile to the modulated groove wall, and the waveform to compensate the error is constructed with delay modulation.

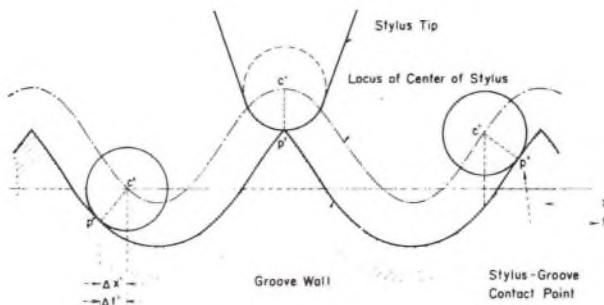


Fig. 2. Spherical stylus tracing a compensated groove wall.

SKEW-COORDINATE THEORY OF COMPENSATING FOR PLAYBACK ERROR

The method is based on the epoch-making discovery of the purely geometrical basis for representing a playback error that at the same time shows how the error may be removed, continuously and perfectly, either as a tracing error or as a tracking-angle error [1].

That tracing error may be described as a curvilinear skew transformation is shown with the help of Fig. 3. The lower part of the figure shows the profile S of a stylus tip engaging the waveform y which modulates the wall of a phonograph groove. The tangential contact is at x . The stylus tip is forced to move in the tangential direction of the arrow shown at the point of contact x . Accordingly the stylus tip, or apex, X on S describes the locus Y embodying the tracing error. In the upper part, S' , y' , and Y' represent slopes of the curves, S , y , and Y , respectively. The curves parallel to S' , corresponding to successive positions of the stylus, thought of as moving from left to right, thus form a skew coordinate grid.

Thus, to form a signal Y in agreement with y , it is

necessary to compensate for the error $X - x$ of the upper part of the figure. This error is given by the inverse function $Cg(y')$ so that the equation of the transformation is

$$X = x - Cg(y'), \quad Y' = y' \quad (2)$$

where y' is the ordinate of the y' curve and C is the vertex radius of curvature of the stylus profile. The curve S' of the figure is for a circular profile for which $g = y'/\sqrt{1 + (y')^2}$, whereas for a parabolic profile, it would be the straight line as the slope function, and this function may be taken even when circular tips are employed, because $(y')^2$ is such a small part of the denominator of g and the major part of the correction is Cy' .

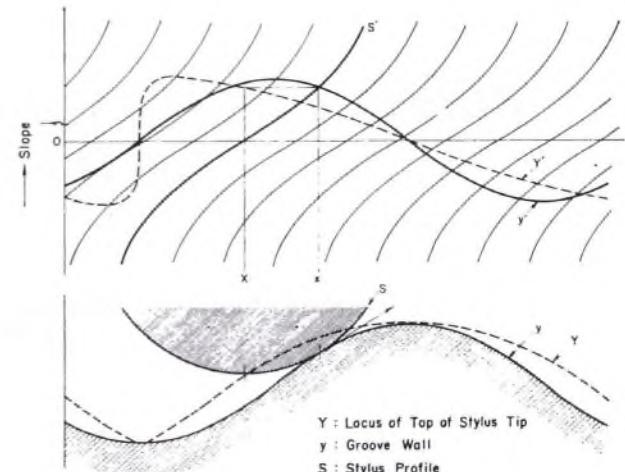


Fig. 3. Skew tracing transformation.

OUTLINE OF THE APPARATUS

In the skew sampling compensator for playback error, the above conception is realized by performing the waveform transformation with the skew-sampling method [2] of Fig. 5, so that the cutting waveform at recording will be the compensated one. The figure illustrates the process of compensation for tracing error in a sinusoidal signal.

The arrangement of the skew-sampling delay modulator is shown in Fig. 4, in which each channel consists of two units, the skew sampler and the boxcar sampler.

First the timing generator and distributor generates the timing pulses of 200 kHz at regular intervals and distributes these cyclically and in sequence among the ten channels so that each receives its distributed timing

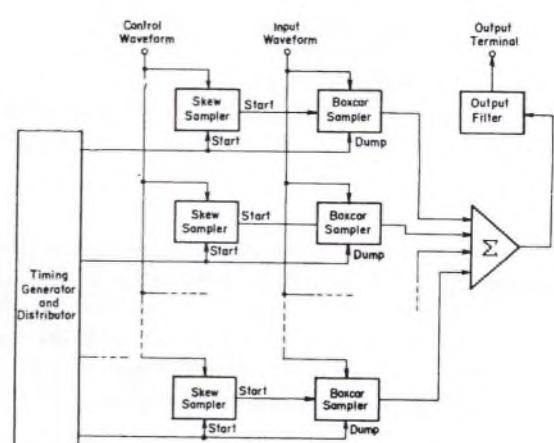


Fig. 4. Skew-sampler flow chart.

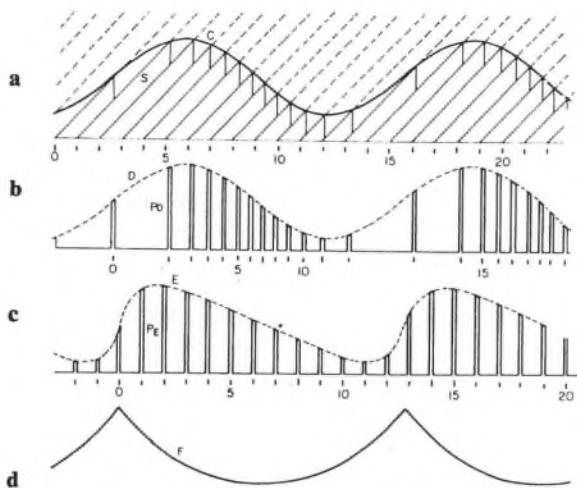


Fig. 5. Waveform and pulses in skew-sampling delay modulator. **a.** Skew sampling of control signal. **b.** Input signal sampled with delay-modulated pulses. **c.** Delayed input pulses to low-pass filter. **d.** Action of cutting stylus (groove waveform).

pulse also at regular intervals of $50 \mu\text{s}$. When the value of the ramp waveform generated in each skew sampler starting at the distributed timing pulse agrees with the control signal value, the ramp is reset to its starting value, and a sampling pulse is emitted to the boxcar sampler. The skew sampler then awaits the next distributed timing pulse to begin its cycle anew.

The boxcar sampler measures the value of the input waveform at the instant it receives the sampling pulse from the skew sampler in a storage unit, such as a capacitor. Upon receiving the next distributed timing pulse, it emits the sample to the common bus conducting the sample to the output filter. It then awaits the next sampling pulse from the skew sampler.

The operation of this system is illustrated in Fig. 5. Fig. 5a shows a control waveform C , or the slope waveform (proportional to a velocity waveform) obtained through differentiation of the original signal, with the ramp waveform S scaled to the stylus profile in the skew sampler. At the times marked by the skew wave attaining the value of the slope waveform, the sampling pulses P_o are caused to occur. The sampling pulses thus modulated in position in accordance with the slope signal

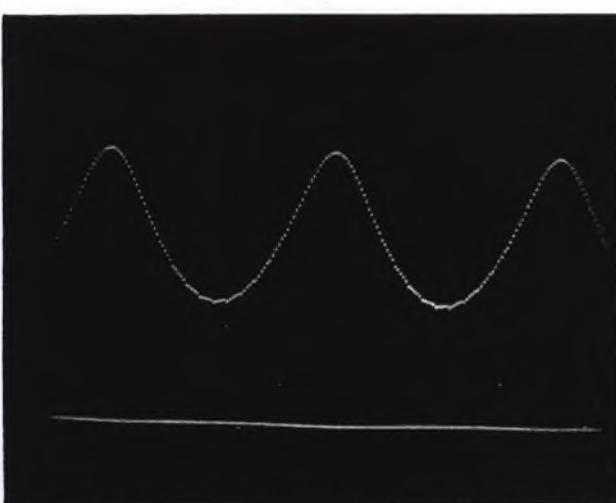


Fig. 6. Oscilloscope display showing synthesis of displacement waveform corrected for tracing error.

are caused to sample the input signal D shown as a similar slope waveform in Fig. 5b, and the resulting sample pulses are held for subsequent emission at equal intervals, with a maximum of $50 \mu\text{s}$ for the holding time of an instantaneous value as shown in Fig. 5c. The envelope waveform E for the emitted pulse train P_E is the waveform compensated for the error of delay modulation and the one supplied to a cutter.

The action of the velocity-controlled cutter is to transform the velocity signal to a displacement signal, the displacement of the cutting stylus, so that the groove waveform in the lacquer master is like that of Fig. 5d. When this displacement waveform F is cut on a disc as sound groove waveform and a pickup stylus tip of spherical form runs along this sound groove, the reproduced velocity waveform becomes almost perfectly like the original waveform D , so that a sound adapted to the original sound will be reproduced without distortion.

The specifications of the manufactured device are as follows.

1) Sampling frequency	200 kHz
2) Delay range for distortion control	$50 \mu\text{s}$
3) Number of channels	10×2
4) Frequency response	$\pm 0.5 \text{ dB}$
5) Dynamic range	20 Hz to 15 kHz more than 90 dB
6) Gain	0 dB.

The instantaneous values of the control signal in compensating for distortion with this device follow the values of the program signal, but modified suitably in level for the stylus radius to be used in playback and for the variation of groove velocity at varying radial positions on the disc, following the relation of Eq. (1).

Other applications within the capability of the compensator are tracking-angle error compensation and compensation during playback, by reversing the control, for uncompensated recordings.

Fig. 6 shows the synthesized output waveform of a 5-kHz signal for tracing compensation delay modulated with this skew sampling device in displacement signal, and Fig. 7 shows that for tracking compensation.

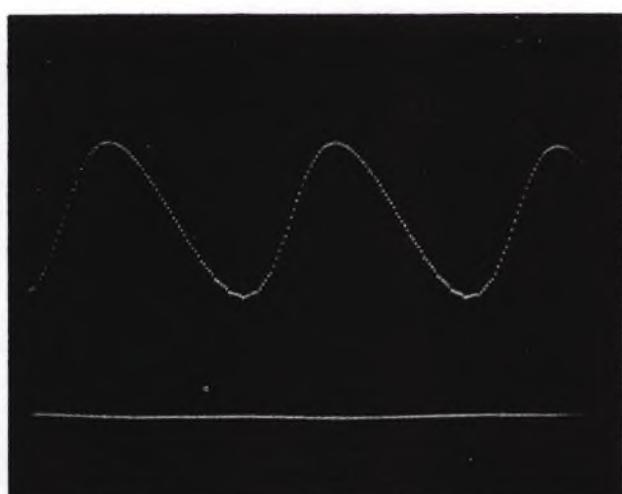


Fig. 7. Oscilloscope display showing synthesis of displacement waveform corrected for tracking error.

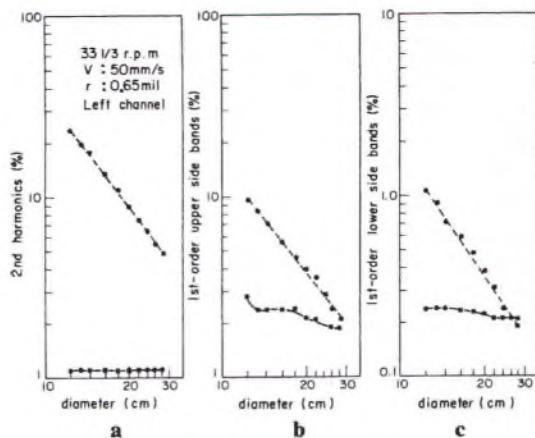


Fig. 8. Measured distortion in velocity signal on a lacquer disc. Upper curves (dashed line)—distortion for uncompensated disc; lower curves (solid line)—distortion for compensated disc. a. Second-order harmonics of 4-kHz signal. b. First-order upper sidebands of 4.0 + 4.4-kHz signal. c. First-order lower sidebands of 4.0 + 4.4-kHz signal. The percent distortion shown in c would be ten times greater in the displacement signal.

EXPERIMENTAL RESULTS

Examples of experimental results, consisting of distortion measurements in the playback of a compensated record in comparison with an uncompensated one, are shown in Fig. 8 [3]. These data were taken at 33⅓ rev/min for a recorded velocity of 50 mm/s (45° direction) using the DENON DL-103 pickup cartridge (stylus radius of 0.65 mil).

For the pure tone signal of 4000 Hz the second-harmonic distortion increases remarkably toward the inner grooves of the disc in the uncompensated case, but this increase does not appear for the skew-sampler compensated case as shown in Fig. 8a. Also, the distortion on the inside of the compensated disc is smaller than on the outside of the uncompensated disc; so the skew-sampler compensated disc is capable of playing back sound of higher quality from the outside to the inside on a disc. By using this compensator, a program source of wide dynamic range will be reproduced through the sound groove essentially without distortion, and "muddiness" of sound caused by the distortion may be removed almost perfectly.

Examples of experimental results obtained in intermodulation distortion measurements are shown in Fig. 8b and c. To obtain these data an intermodulation test signal consisting of 4000 + 4400 Hz in a 1:1 ratio was recorded in the left channel of a disc at various diameters with and without the skew-sampling compensator. The amplitude of the intermodulation upper sidebands, 8400-Hz components to give the percentage values as plotted in Fig. 8b, was measured by means of a wave analyzer in the playback of the velocity-type signal. Next the amplitude of the intermodulation lower sidebands, 400 Hz to give the percentage values as plotted in Fig. 8c, was measured by the same means.

In these results it is clear that intermodulation distortion as well as a harmonic distortion can be reduced by using the skew-sampling compensator.

Although a tracing error is more strongly generated at higher frequencies, the distortion caused by the error can appear at low frequencies as shown in Fig. 8c. Thus it may contribute to poor sound separation, the vague-

ness of localization or the muddiness of sound at the inside of a disc in the absence of compensation. However, these faults may be sufficiently reduced with the skew-sampling compensator.

Also, a sufficient experimental demonstration of the compensation of a tracking-angle error and the compensation of playback distortion from an uncompensated recording has been obtained with a change of the control process of the skew-sampling compensator.

The compensated disc, manufactured by the Nippon Columbia Company, Ltd. as the Master Sonic record, is designed for playback with a spherical stylus of radius 0.65 mil through skew-sampling precompensation. It is possible to play such a disc with an elliptical stylus, of course, and Fig. 9 shows experimental results comparing compensated and uncompensated discs in playback with an elliptical stylus. These data are playback measurements of the second harmonic of 5000 Hz recorded at 35 mm/s as a pure tone and played with an elliptical stylus with radii 0.3 by 0.7 mil, where the compensated disc was designed for 0.65-mil playback. Despite the mismatch in the compensation, the compensated disc produces the better results, showing the insufficiency of using the elliptical stylus as a means of reducing reproduction distortion.

Although the elliptical stylus has been a popular approach to the problem of reducing tracing error, it not only fails to approach perfection in this regard, as the data of Fig. 9 show, but it also causes further problems such as increasing the deformation of the record groove, causing permanent injury and wear, and introducing deformation distortion of the waveform.

The composite distortion will be the less, as these results show, if compensation provides the means for avoiding the use of the elliptical stylus. Thus, indirectly, tracing-error compensation may be regarded as a means for reducing deformation distortion. At frequencies for which the stylus reaction is primarily inertial, tracing compensation, by reducing the acceleration to be supplied by the wall of the groove, can also serve as a direct means for reducing deformation distortion.

Thus, it is not merely a question of reducing tracing distortion, but one of reducing the composite distortion,

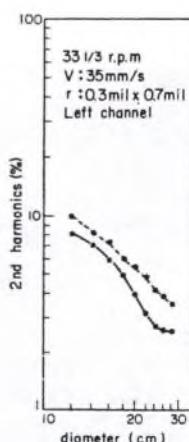


Fig. 9. Distortion measured with an elliptical stylus. Except for the lower recording velocity and the use of the elliptical stylus, the data are for the same conditions as in Fig. 8a. Upper curve (dashed line)—uncompensated disc; lower curve (solid line)—compensated disc designed for 0.65-mil stylus radius.

and of increasing the life of the stylus-disc system. The advantages here lie with spherical-stylus compensation and spherical-stylus playback.

LISTENING QUALITY

As a result of the capability of reproducing the original sound without distortion by the tracing of a spherical stylus tip along the compensated sound-groove waveform cut on a disc as mentioned above, we are able to obtain the following improvements with regard to the quality of the playback sound. 1) The separation and localization of sound of each musical instrument becomes clear, so that the stereophonic sound can be felt more vividly. 2) Particularly the range of high tones at the inside of the disc is remarkably improved compared to the uncompensated disc records. 3) The dynamic range of cutting that had hitherto been suppressed in order to avoid the distortion coming from the difficulty of tracing because of large amplitudes, can be taken larger so that it is possible to enjoy a disc with clear sound. 4) By removing the intermodulation distortion it becomes possible to prevent the generation of compound sounds not existing in the original sound that is produced as mutual beats between sounds of some musical instruments.

CONCLUSION

The experiments described have proven that the skew-sampling method is very effective in compensating for disc-record distortion, and that this method will make possible the manufacture of disc records with a high-

fidelity quality beyond the usual experience.

Tracing compensation is effective in reduction of deformation error, too, and the listening effect will be remarkable with the upgrading of the playback system. We are entering the multichannel stereophonic age and just starting to realize the multichannel reproduction system by using disc records.

So the distortion of disc reproduction will be an even more important factor to improve listening quality in multichannel systems than in the usual two-channel stereo system.

In the future the distortion compensation of the disc record will be absolutely necessary with the advance of audio system from recording to reproducing and with the complicated transmission signal.

ACKNOWLEDGMENT

The authors wish to thank Dr. D. H. Cooper of the University of Illinois for his continuing guidance and helpful advice in the development of the skew-sampling device and the experiments performed with it.

REFERENCES

- [1] D. H. Cooper, "Integrated Treatment of Tracing and Tracking Error," *J. Audio Eng. Soc.*, vol. 12, p. 2 (1964).
- [2] D. H. Cooper, "Continuous Delay Regulator for Controlling Recording Error," *J. Audio Eng. Soc.*, vol. 14, p. 7 (1966).
- [3] S. Washizawa, *J. Acoust. Soc. Jap.*, vol. 28, p. 3 (1972).

THE AUTHORS

Shigetaka Washizawa was born in Tokyo, Japan, in 1938. He received a degree in electrical engineering from the Waseda University in 1961. Following graduation he joined the Nippon Columbia Company, Ltd., where he is doing research and development in the field of electroacoustics, especially on stereo sound fields, room acoustics, speech recognition, phonograph recording, and four-channel stereo systems.

Mr. Washizawa is a member of the Acoustical Society of Japan.

Tomofumi Nakatani was born in Hokkaido, Japan, in 1943. He graduated from Shizuoka University, Hamamatsu, Japan, in 1968. Since 1968 he has been with the Research Laboratories of Nippon Columbia Company, Ltd., as an Acoustic Engineer. He has been

engaged in research and development of problems in the field of electroacoustics, especially on tracking-error compensation and four-channel stereo systems.

Takeo Shiga was born in Tokyo, Japan, in 1924. He received a degree in engineering physics in 1946 and a doctorate in applied physics in 1962, both from Tokyo University. In 1946 he joined Nippon Columbia Company, Ltd., where he has been responsible for the development of loudspeakers, phonograph pickups, and stereo phonographs. In recent years his studies have been devoted to stereophonic sound. At present he is the General Manager of the Acoustical Products and Musical Instruments Division.

Dr. Shiga is a member of the Audio Engineering Society and the Acoustical Society of Japan.

A Theory of Scanning Loss in Phonographs*

JAMES V. WHITE

Stevens Institute of Technology, Hoboken, New Jersey 07030

Scanning loss is defined as that component of the total mechanical playback loss which is independent of the driving-point impedance of the stylus. A phenomenological theory of this loss is developed to account for the main features in scanning-loss data obtained by J. G. Woodward. The theory in its simplest form states that the magnitude of the scanning loss as a function of recorded wavelength depends on three parameters: (1) the width w_c of the stylus-groove contact region, which determines the frequencies of the dips in scanning-loss data; (2) an asymmetry parameter a that measures a component of asymmetry in the stylus-groove interaction, which determines the depths of these dips; and (3) an envelope frequency f_e that measures the extent to which the groove deformation permits approaching parts of the groove modulation to influence the stylus motion before this modulation actually touches the stylus, which determines the envelope of scanning-loss data. Direct verification of the theory will be possible when measurements of the phase responses of phonographs become available.

INTRODUCTION: Mechanical playback losses in phonographs can be classified in terms of two categories: those losses that are dependent on the driving-point impedance of the playback stylus and those that are independent. The type of loss that is independent of the stylus impedance is known as scanning loss and forms the subject of this paper. This loss can persist even if we use ideal pickups having vanishingly-small stylus impedances.

A common result of scanning loss is the decrease in high-frequency response that occurs as the pickup translates toward the center of the record. This decrease is known as translation loss¹ or diameter loss [1], [2], [8]

and varies anywhere from 3 dB to 15 dB at 15 kHz in commercial phonographs.

Kornei [3] gave the first theoretical treatment of mechanical playback losses in 1941. He assumed for mathematical convenience that the groove deformation caused by the stylus can be described by using Hertz's theory [4], [5], [6] for static contact between perfectly elastic bodies. Nine years later Miller [7] corrected a serious error in Kornei's analysis, extended the calculations, provided new experimental results, and also introduced some new ideas in the theoretical treatment, such as his "scanning-loss function," which was intended to extend the validity of the analysis to shorter wavelengths.

Since 1950 additional scanning-loss data have been presented by Kantrowitz [8] and Bastiaans [9]. Kantrowitz's data are limited to frequencies that lie below the first critical scanning-loss frequency, and they show that Miller's theory is useful within this lower frequency range. Bastiaans' data also show good agreement with Miller's theory in this frequency range. In the vicinity of the first critical scanning-loss frequency, however, the

* A preliminary version of this paper was presented October 13, 1970 at the 39th Convention of the Audio Engineering Society, New York.

¹ The term translation loss has been used by some authors to denote one component of the scanning loss defined in this paper. This use of the term originated in Miller's analysis [7].

agreement is not very satisfying. Data presented more recently by J. G. Woodward, in contrast, extend well beyond this first critical frequency and show that scanning loss in this higher frequency range is completely different from the loss predicted by Miller's theory. This means, for instance, that Miller's theory should not be trusted to predict the frequency response in high-density recording where scanning loss is the major playback loss, and the frequencies extend beyond the first critical scanning-loss frequency.

This paper presents some of Woodward's data on scanning loss and describes a theory of scanning loss that takes into account three main features of these data.

ASSUMPTIONS AND NOTATION

We shall take as our starting point the linear mathematical model of the stylus-groove interaction that was developed in reference [10]. This model applies to vertically modulated records and is based on six assumptions: (1) surface noise and inhomogeneities in the deformation properties of the record plastic can be ignored; (2) the stylus-groove interaction at the left groove wall is a mirror image of the one at the right wall; (3) the stylus has a zero-degree vertical tracking angle and performs vertical oscillations without rotation in response to vertical groove modulation; (4) the stylus motion is completely unresponsive to both the horizontal friction force and the moments acting on the stylus; (5) the stylus dynamics are already known in the form of the driving-point impedance of the stylus; and (6) the groove modulation and the vibrational displacements of the stylus are both small enough that the time-varying component of force acting on the stylus (because of groove deformation) depends on them linearly.

This mathematical model contains three functions of time: the vertical force $F(t)$ acting on the stylus due to groove deformation; the resulting vertical displacement $q(t)$ of the stylus; and the modulation function $M(t)$ that denotes the vertical displacement of the groove bottom (due to groove modulation) at a location directly beneath the stylus at instant t . These single-valued functions of time, $-\infty < t < \infty$, are each assumed to be positive in the upward direction as shown in Fig. 1. The modulation function $M(t)$ is assumed to have zero average value, and in an unmodulated groove $M(t) = 0$. The deformation force $F(t)$ is split into its average value F_0 and its variational part $F_v(t)$: $F(t) = F_0 + F_v(t)$. The average value F_0 is a constant equal to the tracking force referred to the stylus tip. In the same way, the stylus displacement function is written as $q(t) = q_0 + q_v(t)$. This function is defined so that $q(t) = 0$ if the stylus traces an unmodulated groove without deformation. The mean value q_0 , which represents the mean groove deformation, is therefore a negative number because the average groove indentation corresponds to a downward average stylus displacement.

Fourier transforms are distinguished from time functions by overbars; e.g., the transform of the variational force function is defined (in the sense of generalized functions [11] if necessary) by

$$\bar{F}_v(f) = \int_{-\infty}^{\infty} F_v(t) \exp(-j2\pi ft) dt. \quad (1)$$

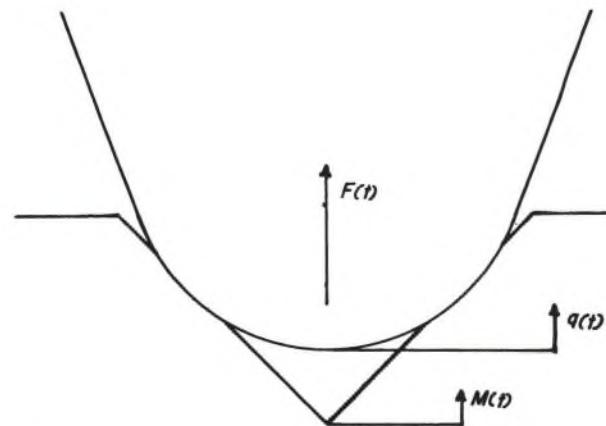


Fig. 1. Cross section of stylus and groove in vertical plane that passes through center of stylus and is oriented normal to groove axis. Shown are vertical force $F(t)$ due to groove deformation, vertical position $M(t)$ of groove bottom, and vertical position $q(t)$ of stylus.

FREQUENCY RESPONSE FUNCTION

The mechanical frequency-response function $\bar{R}(f)$ of a phonograph pickup is defined² as that transfer function which relates the modulation spectrum $\bar{M}(f)$ to the resulting variational displacement spectrum $\bar{q}_v(f)$ of the stylus:

$$\bar{q}_v = \bar{R}\bar{M}. \quad (2)$$

In reference [10] it is shown that the response function $\bar{R}(f)$ can always be written, under the assumptions given above, as

$$\bar{R} = \bar{P}\bar{S}. \quad (3)$$

The frequency-response function is shown here factored into two parts: $\bar{S}(f)$ is called the scanning function and should not be confused with Miller's scanning-loss function [7]; while $\bar{P}(f)$ is called the pickup function.

The pickup function $\bar{P}(f)$ does not concern us in this paper because it depends on the driving-point impedance of the stylus. As long as this stylus impedance is much smaller than the groove impedance, the pickup is free of mechanical pickup loss because $\bar{P}(f) \approx 1$, and the pickup is free of mechanical resonances. Under this condition the response function $\bar{R}(f)$ is determined by the scanning function $\bar{S}(f)$, and any deviations in frequency response are due entirely to scanning loss. The magnitude of scanning loss, in decibels, is given by

$$\text{scanning loss (dB)} = 20 \log |\bar{S}(f)|. \quad (4)$$

² The frequency-response function for the "pressure-tracing" phonograph system introduced in [14] is defined differently because the pickup in this system measures deformation force rather than stylus displacement or velocity. Thus the response function for "pressure tracing," denoted $\bar{R}_0(f)$, relates the modulation to the resulting variational deformation force: $\bar{F}_v = \bar{R}_0 \bar{M}$. Reference [10] shows that with an ideal pickup $\bar{R}_0 = j2\pi f \bar{Z}_0 \bar{S}$, where $\bar{Z}_0(f)$ is the groove impedance referred to the stylus, and $\bar{S}(f)$ is the scanning function.

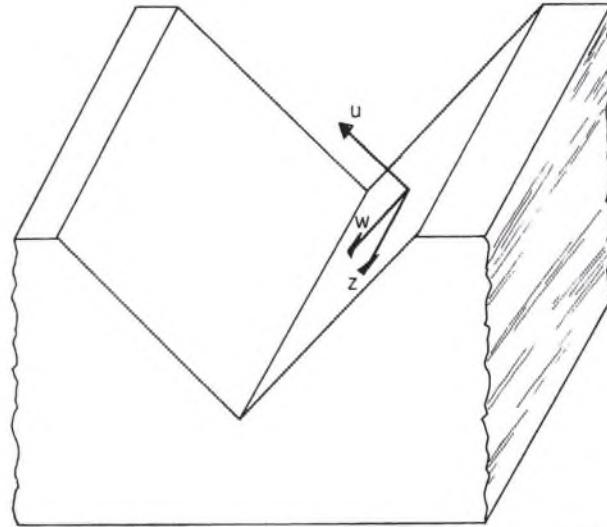


Fig. 2. Rectangular u - w - z coordinate system. u passing through the center of curvature of the stylus (not shown) and normal to unmodulated groove wall, z parallel with the groove axis and tangent to the unmodulated wall.

As stated in [10], $\bar{S}(0) = 1$, so that scanning loss is always zero at sufficiently low frequencies.

PHYSICAL BASIS OF SCANNING LOSS

The physical basis of scanning loss can be seen by considering what happens if the stylus is clamped so that it can not vibrate while tracing a slightly modulated groove. Figure 2 shows the coordinate system we shall use. Figure 3 shows a clamped motionless stylus at instant t (as seen when viewed along the w axis) in sliding contact with the right groove wall. The point $z = 0$ is directly beneath the stylus. The groove wall is moving to the left along the positive z axis with velocity v , and sliding friction is assumed to cause the pileup of plastic on the right or leading side of the stylus. The width of the stylus-groove contact extends from $z = z_1$ to $z = z_2$. In the case shown here this width includes more than one complete wavelength of the modulation. This state of groove deformation is plausible, although I have not had an opportunity to observe if it actually occurs to the scale shown here. The $1/\sqrt{2}$ factor multiplies the modulation function because the groove wall is inclined 45° from the vertical.

The contact width w_c fluctuates only slightly if the stylus is held motionless because the modulation is as-

sumed to be small compared to the average indentation. The net force acting on the stylus contains a small fluctuating component due to the modulation. This fluctuation component would normally cause the stylus to vibrate, but in the present mode of operation the stylus is being held motionless. The fluctuating component of force depends on the modulation, not just at a single point under the stylus, but throughout the region of contact. In fact, because of the pileup the fluctuating force will depend to some extent on the modulation that lies to the right of the stylus even though this modulation has not yet moved within the contact region proper. The force may also depend on the modulation in the region to the left of the trailing edge.

The cause of scanning loss can now be explained. The low hills and shallow valleys of the modulated groove produce a net fluctuating force on the stylus. The magnitude of these fluctuations decreases as the wavelength decreases because the positive force contribution from the hills is progressively canceled by the negative contribution from the valleys. If the interaction were spacially symmetric, then we should expect perfect cancellation between these positive and negative force contributions when just the right number of recorded wavelengths fit inside the contact region. Scanning-loss data show, however, that perfect cancellation does not occur in practice. Instead, the scanning loss alternately goes through relative maxima and minima as the recorded frequency is increased (wavelength is decreased).

MEASURED SCANNING LOSS

Two beautiful examples of scanning loss are shown in Fig. 4. Although these experimental results were previously discussed in [12], they will be given another detailed discussion here because they provide a foundation for part of the theory that is developed later in this paper. These data were kindly provided by J. G. Woodward of RCA, Princeton, N.J. Dr. Woodward obtained these frequency response curves by playing special short-wavelength test records at the very low groove speeds of 2.2 and 4.4 cm/s. The pickup and tone-arm he used were free of resonances in the frequency range shown here so that these curves are believed to represent variations in the magnitude of the scanning function $\bar{S}(f)$, rather than variations in the pickup function $\bar{P}(f)$. The velocity-sensitive pickup used was fitted with a stylus having a 0.3 mil tip radius, the grooves were vertically modulated at 0.3 cm/s, and the tracking force was 1 g.

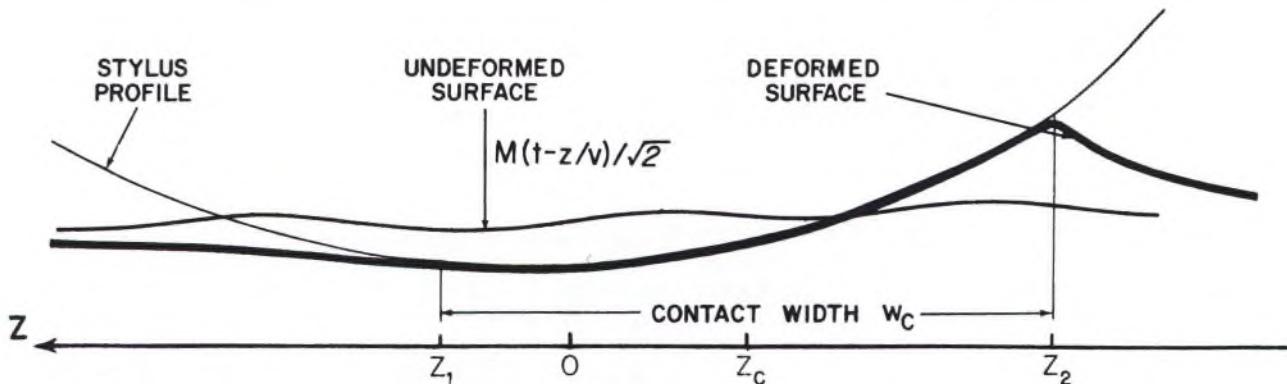


Fig. 3. Cross section in u - z plane of the stylus deforming the slightly modulated groove wall.

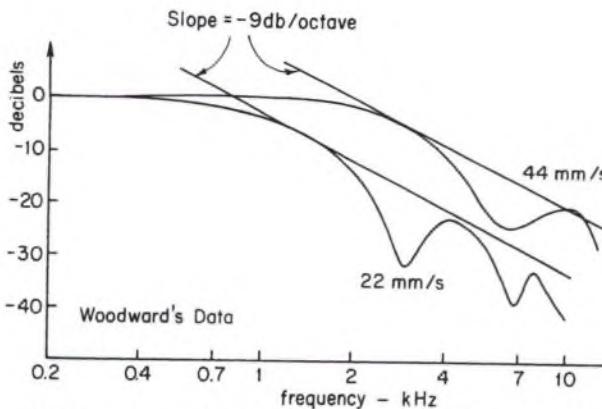


Fig. 4. Measured scanning loss versus frequency after Woodward.

There are three primary features of these data that any reasonably complete theory should account for: the locations of the dips; the depths of these dips; and the slope of the envelope. With many high quality commercial phonographs operating at 33.3 r/min, the first dip is located at a critical frequency above 50 kHz when the outer grooves of a 12-in record are being played, whereas this dip drops to about one-half this frequency at the inner grooves. A good rule of thumb is that the critical frequency is directly proportional to the linear groove velocity.

According to Dr. Woodward, these curves are rather typical of many that were measured under a variety of conditions. The dips are not always as clearly defined as in this example and sometimes do not show up at all. A reduction in tracking force is the most effective means of smearing out the dips, but other not altogether predictable factors may do it also. In addition, Woodward noted that the near proportionality of the dip frequencies to the groove speed indicates a wavelength dependence. Some of his results contradicted Miller's theory. While they showed that the frequency of the first dip does decrease with increasing tracking force, the trend was erratic and the shift not very great, perhaps a 20% decrease in frequency for a trebling of the force, which may be compared to Miller's prediction of a 30% decrease. Above the frequency of the first dip these data do not even bear a qualitative resemblance to Miller's predictions. Woodward found that the dips definitely move to lower frequencies when the stylus-tip radius is increased. Very similar results were obtained for both vertical and lateral modulation.

SCANNING THEORY

Friction and visco-elastic-plastic deformation effects complicate the stylus-groove interaction to such an extent that it seems impractical to build up a scanning theory from the fundamental principles of continuum mechanics. Instead we shall develop a phenomenological description of scanning loss by playing a curve-fitting game with Woodward's data. At first glance this may appear to be a dangerous and arbitrary venture, promising little real insight into the physics of scanning loss. I feel, however, that the results obtained here do not support this view, but instead bring to light some engaging questions about scanning loss and a plausible physical picture of its most important details. Several questions must remain unanswered, however, until the phase re-

sponses of phonograph pickups are measured at sufficiently short wavelengths.

Our goal is to develop a mathematical expression for the scanning function $\bar{S}(f)$. This expression must have a clear and reasonable physical interpretation and must also yield calculated scanning-loss that is consistent with Woodward's data with respect to the locations of the scanning dips, the depths of these dips, and the -9 dB/octave slope of the high-frequency (short wavelength) envelope.

We start by considering the locations of the dips; Fig. 4 shows that the frequency of the second dip is about 2.2 times the frequency of the first dip. By looking through a pictorial dictionary of Fourier transforms [13] one discovers a remarkable fact: our friend the zero-order Bessel function of the first kind, J_0 , has a lobed structure like that of Woodward's data. If we take $J_0(c_1 f)$ (c_1 = constant to be determined later) as our first guess for the scanning function and denote this guess by $\bar{S}_1(f)$, then the calculated scanning loss shown in Fig. 5 is obtained. Although the slope of the envelope and the depths of the dips do not match Woodward's data, the relative locations of the first two dips do match closely: the ratio of the second dip frequency to the first dip frequency is 2.3, which differs from the ratio in Woodward's data by only 4%.

To see how the first trial scanning function $\bar{S}_1(f)$ can be altered to yield a better match with Woodward's data, we shall reconsider the variational displacement $q_v(t)$ of the stylus at instant t . According to Eqs. (2) and (3) the spectrum $\bar{q}_v(f)$ can be written as follows, provided that $\bar{P}(f) = 1$ so that the pickup loss can be neglected:

$$\bar{q}_v = \bar{S}\bar{M}. \quad (5)$$

The expression we need for $q_v(t)$ is obtained by taking the inverse transform of Eq. (5), which yields the following convolution integral:

$$q_v(t) = \int_{-\infty}^{\infty} S(\tau) M(t-\tau) d\tau. \quad (6)$$

To give a physical interpretation to this equation we change the dummy variable of integration: $\tau = z/v$, $d\tau = dz/v$, where v = linear groove velocity and z corresponds to the z axis in Fig. 3. Equation (6) then reads

$$q_v(t) = \int_{-\infty}^{\infty} S(z/v) M(t-z/v) dz/v, \quad (7)$$

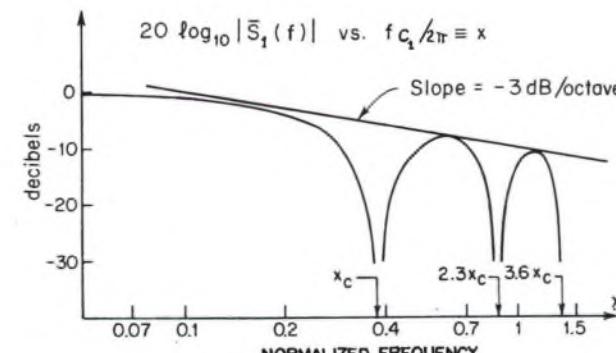


Fig. 5. Scanning loss of scanning function $\bar{S}_1(f) = J_0(c_1 f)$.

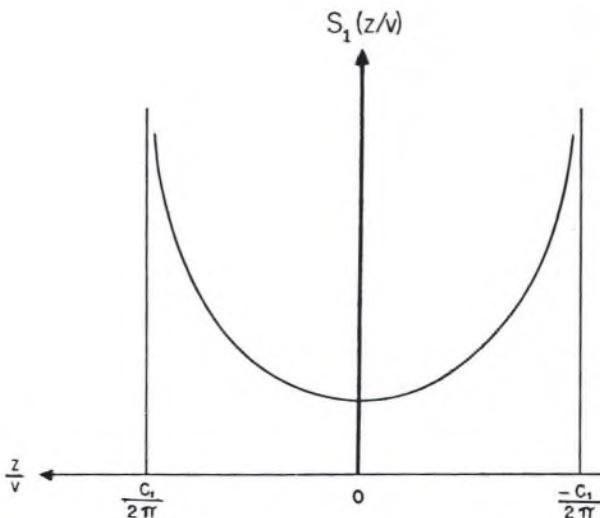


Fig. 6. Schematic drawing of the weighting function $S_1(z/v)$.

which, for future convenience, can be represented more compactly by using the usual asterisk notation for convolution:

$$q_v(t) = S^* M(t). \quad (8)$$

In Eq. (7) $M(t - z/v)$ corresponds to the modulation shown in Fig. 3, and the function $S(z/v)$ (inverse transform of scanning function $\bar{S}(f)$) plays the role of a weighting function; the integral can be interpreted as a weighted spacial average of the modulation. We shall therefore refer to $S(z/v)$ as the weighting function. At any arbitrary instant t , this weighting function specifies the relative importances that different parts of the groove wall have in determining the stylus displacement $q_v(t)$ at that same instant.

A number of things can be learned about $S(z/v)$ from Fig. 3. For instance, the displacement of the stylus clearly depends on that part of the modulation which is contained within the region of physical stylus-groove contact, $z_1 \leqq z \leqq z_2$. In consequence we know that $S(z/v)$ should not vanish in this range of z . Also the stylus displacement depends to some extent on the modulation that lies just in front of the stylus in the pileup region as well as just behind in the trailing region. Thus we expect $S(z/v)$ to be non-zero in a certain neighborhood of the physical contact region. On the other hand, the stylus displacement at instant t is clearly not influenced by modulation that is distant from the contact region at that instant. Thus we must require that $S(z/v) \rightarrow 0$ as $|z| \rightarrow \infty$.

Two other characteristics of $S(z/v)$ can be deduced without reference to Fig. 3. These are (1) that $S(z/v)$ must be a real function because $q_v(t)$ and $M(t)$ are both real, and (2) that the total area under the graph of $S(z/v)$ must be unity; i.e.,

$$\int_{-\infty}^{\infty} S(z/v) dz/v = 1. \quad (9)$$

This normalizing condition on $S(z/v)$ is a direct result of the fact that $\bar{S}(0) = 1$.

We are now in a position to assess the physical reasonableness of our first guess $\bar{S}_1(f) = J_0(c_1 f)$ for the scanning function. We previously set $\bar{S}_1(f) = J_0(c_1 f)$, where c_1 is a positive

scaling parameter that is to be determined. (Since $J_0(0) = 1$, this equation clearly satisfies the requirement that $\bar{S}_1(0) = 1$.) Taking the inverse transform yields the following expression for the weighting function:

$$S_1(z/v) = \frac{2 \operatorname{rect}(2\pi z/v c_1)}{c_1 \sqrt{1 - (2\pi z/v c_1)^2}} \quad (10)$$

where the rectangle function $\operatorname{rect}(x) = 1$ for $|x| < 1$, and $\operatorname{rect}(x) = 0$ for $|x| > 1$. To determine parameter c_1 we consider the graph of $S_1(z/v)$ shown in Fig. 6. This weighting function does satisfy the requirement that $S_1(z/v) \rightarrow 0$ as $|z| \rightarrow \infty$; in fact, $S_1(z/v) = 0$ for all $|z| > vc_1/2\pi$. The perfect symmetry about $z = 0$ is unphysical, however, because Fig. 3 clearly implies that the stylus-groove interaction is spatially unsymmetrical. Furthermore, Fig. 3 shows that the physical contact region is not centered at $z = 0$. A physically reasonable means for choosing parameter c_1 is to choose it so that the width of weighting function $S_1(z/v)$ equals the width, $w_c = z_1 - z_2$, of the physical contact region. Thus we set $vc_1/\pi = w_c$, which yields

$$c_1 = \pi w_c / v. \quad (11)$$

We can also account for the fact that the center of the contact region is at $z = z_c$ (rather than $z = 0$) by shifting the weighting function $S_1(z/v)$ appropriately. This is easily done by substituting $z - z_c$ for z in Eq. (10). We then obtain, with the help of Eq. (11), the following new weighting function $S_2(z/v)$:

$$S_2(z/v) = \frac{2 v \operatorname{rect}(x)}{w_c \sqrt{1 - x^2}}, \quad x \equiv 2(z - z_c)/w_c. \quad (12)$$

The corresponding scanning function is

$$\bar{S}_2(f) = J_0(\pi w_c f / v) \exp(-j2\pi f z_c / v). \quad (13)$$

Our shifting the weighting function a distance z_c has introduced the exponential factor, which merely produces a phase shift proportional to frequency. The scanning loss has not been altered, however, because $|\bar{S}_2(f)| \equiv |\bar{S}_1(f)|$.

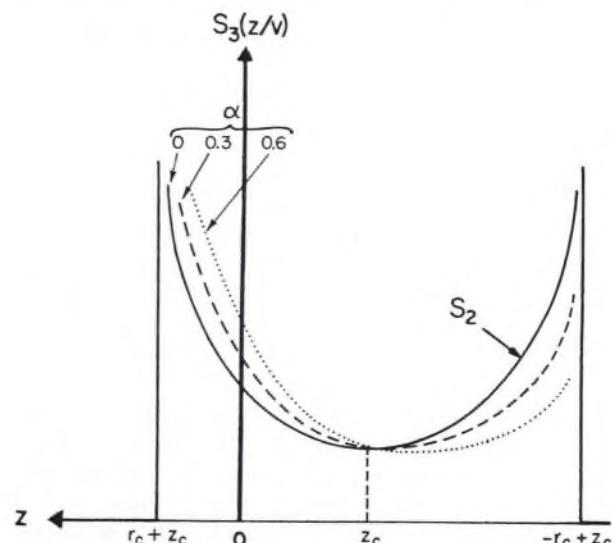
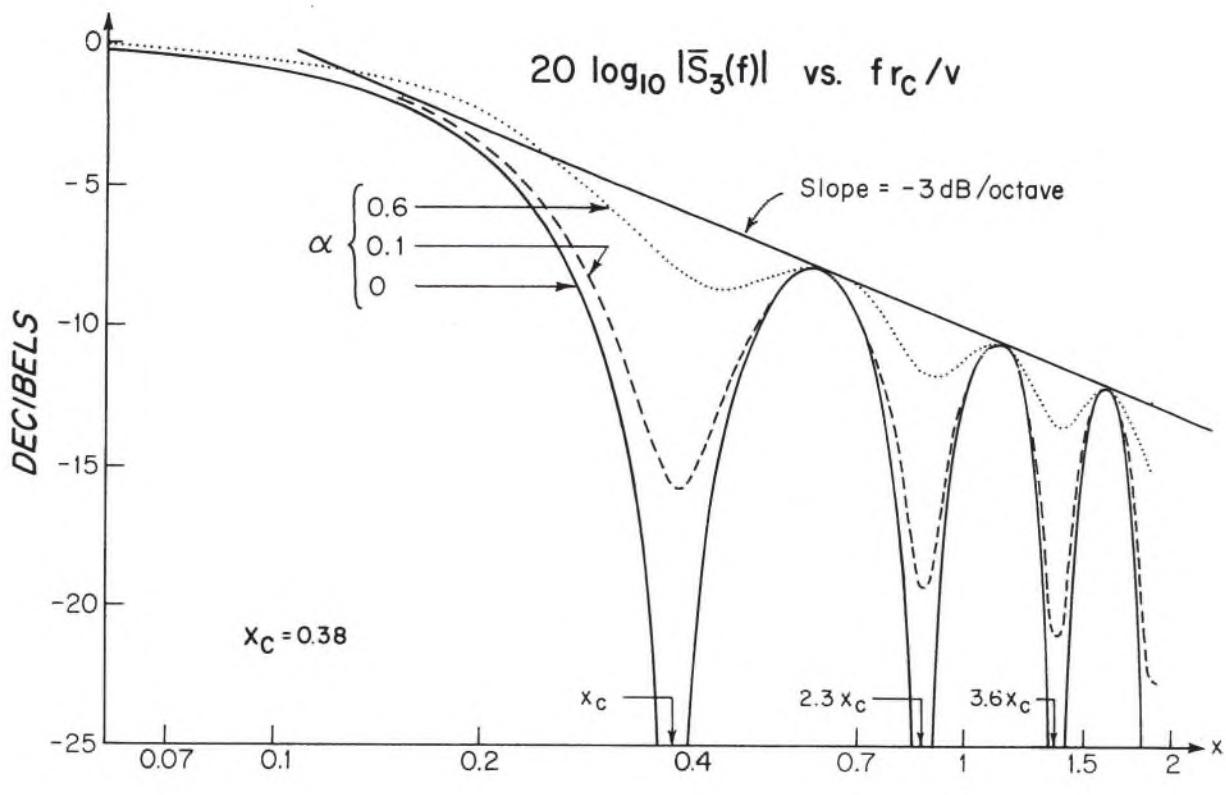


Fig. 7. Schematic drawing of the weighting function $S_3(z/v)$. $r_c \equiv w_c/2$.



NORMALIZED FREQUENCY ($X = f r_c / v$)

Fig. 8. Scanning loss of scanning function $\bar{S}_3(f)$. $r_c \equiv w_c/2$.

The scanning function $\bar{S}_2(f)$ can be written more compactly in terms of the wavelength $\lambda = v/f$:

$$\bar{S}_2(v/\lambda) = J_0(\pi w_c/\lambda) \exp(-j2\pi z_c/\lambda). \quad (14)$$

$\bar{S}_2(f)$ predicts that the first scanning dip occurs at the critical frequency $f_c = 0.76v/w_c$, which corresponds to the critical wavelength $\lambda_c = 1.3w_c$; the second dip frequency is $2.3f_c$ and the third is $3.6f_c$. The spacing between the first and second dips agrees with Woodward's data to within 4% as stated previously, and the relation between the critical wavelength λ_c and the contact width w_c appears to be in agreement with existing data, although no direct measurements of w_c have been reported, and so only estimates of this contact width can be made.

It is a curious fact that the first critical frequency of $\bar{S}_2(f)$ turns out to be only 3% higher than the critical frequency predicted by Miller's "translation-loss function" (which he, Kantrowitz, and Bastiaans all used for interpreting most of their data on scanning loss) if the contact width w_c is set equal to the width predicted by Hertz's theory of elastic contact.

DEPTHS OF SCANNING DIPS

Having obtained a scanning function $\bar{S}_2(f)$ that adequately accounts for the locations of the scanning dips, we now show how an improved estimate of the scanning function can be constructed that also adequately accounts for the depths of these dips. We assert that the depths of the dips are controlled for the most part by asymmetries in the weighting function $S(z/v)$. In defense of this assertion we note that a more physical asymmetric

weighting function $S_3(z/v)$ can be formed from $S_2(z/v)$ by use of the asymmetric linear function $L(z/v)$:

$$S_3(z/v) \equiv L(z/v) S_2(z/v), \quad (15)$$

where

$$L(z/v) \equiv 1 - ax, \quad x \equiv 2(z - z_c)/w_c, \quad (16)$$

in which the real parameter a controls the degree of asymmetry as shown in Fig. 7. Since the resulting new scanning function is the convolution $\bar{S}_3(f) = \bar{L}^* \bar{S}_2(f)$, and since $\bar{L}(f) = (1 + 2az_c/w_c) \delta(f) - jav/(\pi w_c) \delta'(f)$, where $\delta(f)$ and $\delta'(f)$ are the delta function and its derivative, it follows that

$$\begin{aligned} \bar{S}_3(f) = & \\ & [J_0(\pi fw_c/v) + jaJ_1(\pi fw_c/v)] \exp(-j2\pi fz_c/v), \end{aligned} \quad (17)$$

where J_1 is the Bessel function of order one.

In contrast to the previous scanning functions, $\bar{S}_3(f)$ is complex even if the contact area is centered under the stylus so that $z_c = 0$. Since the real and imaginary parts of $\bar{S}_3(f)$ do not vanish together, its modulus never vanishes for $a \neq 0$, and consequently the dips are not infinitely deep. $\bar{S}_3(f)$ is plotted in Fig. 8 for different values of the asymmetry parameter.

SLOPE OF THE ENVELOPE

The slope of the envelope for scanning function $\bar{S}_3(f)$ is -3dB/octave , whereas Woodward's data show a slope of roughly -9dB/octave . This discrepancy arises because of the infinities at the boundaries of the weighting function $S_3(z/v)$. To show this we form a new scanning func-

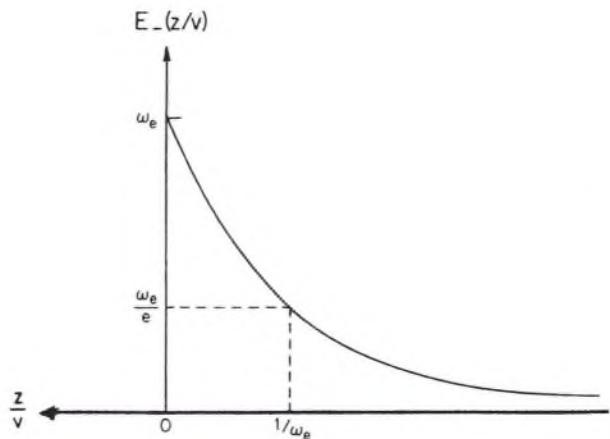


Fig. 9. Smoothing function $E_-(z/v)$. $\omega_e \equiv 2\pi f_e$.

tion $\bar{S}_4(f)$, which we force to have the correct envelope by multiplying $\bar{S}_3(f)$ by a suitable envelope correction function $\bar{E}(f)$:

$$\bar{S}_4(f) \equiv \bar{E}(f)\bar{S}_3(f). \quad (18)$$

For Woodward's data we need a correction function that is unity for low frequencies and falls off at the rate of -6dB/octave at high frequencies. Furthermore, $\bar{E}(f)$ must have a real transform $E(z/v)$ because the weighting function $S_4(z/v)$ must be real. Since scanning-loss data give only the modulus of the scanning function, and since this is all we have to go on, two simple possibilities are the following:

$$\bar{E}_\pm(f) \equiv 1/(1 \pm jf/f_e), \quad (19)$$

where the positive parameter f_e is called the envelope frequency because it is the frequency at which $\bar{E}_\pm(f)$ yields

a -3dB envelope correction, and where either a plus or a minus sign can be used to satisfy our present requirements. The physical significance of plus, as opposed to minus, will be seen shortly. To calculate the new weighting function $S_4(z/v)$, we need the fact that the inverse transform of $\bar{E}_\pm(f)$ is

$$E_\pm(z/v) = 2\pi f_e \exp(\mp 2\pi f_e z/v) U(\pm z/v), \quad (20)$$

where the unit step function $U(x)$ is defined as $U(x) = 1$ for $x > 0$ and $U(x) = 0$ for $x < 0$. This truncated exponential is shown in Fig. 9 for the case of the minus sign. The new weighting function $S_4(z/v)$, which can be written as the following convolution

$$S_4(z/v) = E_\pm * S_3(z/v), \quad (21)$$

is shown in Fig. 10 for the case of the minus sign, asymmetry parameter $\alpha = 0$, envelope frequency $f_e = 850$ Hz, and contact width $w_c = 5.6 \mu\text{m}$. As we shall see later, these numerical values for f_e and w_c make a reasonable fit with Woodward's data.

Both infinities have been removed from $S_3(z/v)$ by this smoothing operation, and $S_4(z/v)$ exponentially weights the modulation in the pileup region in front of the leading edge of the contact region. We call this an example of purely anticipatory scanning because modulation influences the stylus motion before it reaches the contact region proper. In a complimentary way $E_+(z/v)$ describes pure memory scanning because it yields a weighting function that exponentially weights the modulation in the trailing region behind the stylus-groove contact. The choice of a plus sign therefore implies that the scanning operation "remembers" modulation after it has left the physical contact area of width w_c , while the choice of a minus sign implies that the scanning operation "anticipates" the modulation before it touches the stylus.

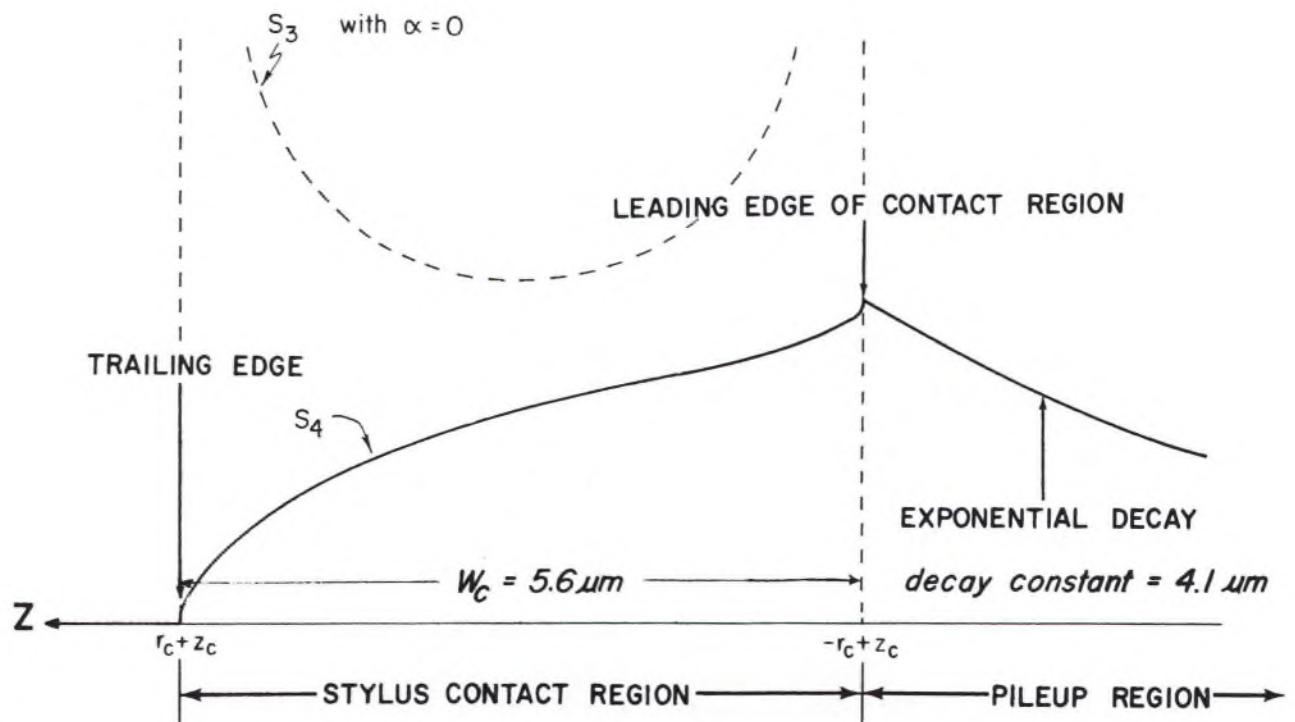


Fig. 10. Weighting function $S_4(z/v)$ with $E = E_-(z/v)$, $f_e = 850$ Hz, $w_c = 5.6 \mu\text{m}$, and $\alpha = 0$. Decay constant = $v/(2\pi f_e)$ = the distance for an exponential decay of 63%.

In other words, the effective scanning aperture extends beyond the physical contact area.

It is unlikely, however, that the scanning aperture extends only in front of or only behind the physical contact area. It undoubtedly does both. Thus a more reasonable envelope correction function $\bar{E}_1(f)$ can be formed as a linear combination of $\bar{E}_+(f)$ and $\bar{E}_-(f)$:

$$\bar{E}_1(f) \equiv C_0 [\bar{E}_+(f) + C\bar{E}_-(f)], \quad (22)$$

where we choose the parameter $C_0 = 1/(1+C)$ so that $\bar{E}_1(0) = 1$. The real parameter C determines the relative strengths of the anticipation and memory terms. By substituting $\bar{E}_\pm(f)$ from Eq. (19) (with $f_e \rightarrow f_+, f_-$) in Eq. (22) we obtain

$$\bar{E}_1(f) = \frac{1 - if/f_c}{(1 + if/f_+) (1 - if/f_-)}, \quad f_c \equiv \frac{f_- f_+ (1+C)}{f_+ - Cf_-}. \quad (23)$$

This result can produce a variety of different scanning-loss characteristics, depending on the relative values of f_- , f_+ , and f_c . For example, if $f_c = f_+$ (which requires that $f_- = f_+/(1+2C)$), then

$$|\bar{E}_1(f)| = [1 + (f/f_-)^2]^{-1/2} \quad (24)$$

which is simply equal to $|\bar{E}_-(f)|$. Similarly, $f_c = -f_-$ (which requires that $f_+ = Cf_-/(2+C)$) implies that

$$|\bar{E}_1(f)| = [1 + (f/f_+)^2]^{-1/2} \quad (25)$$

which is equal to $|\bar{E}_+(f)|$.

These examples show that measurements of the modulus of $\bar{S}_4(f) = \bar{E}_1(f)\bar{S}_3(f)$, as opposed to measurements that include its phase, can not provide enough information for us to distinguish between pure anticipatory scanning, pure memory scanning, or a linear combination of the two. Unfortunately the needed phase responses of phonographs have apparently not been measured.

DISCUSSION

We have constructed a mathematical expression for the scanning function $\bar{S}(f)$ that has a clear physical interpretation and that yields predicted scanning loss that is consistent with the three main features in Woodward's data. Our final form for the scanning function contains the groove speed v and six additional parameters, but only three of these are needed to account for Woodward's data because these data do not include phase information. These are (1) the contact width w_c , which determines the locations of the scanning dips, (2) the asymmetry parameter a , which determines the depths of the dips, and (3) the envelope frequency f_e that determines the slope of the envelope. The three additional parameters are z_c , which determines a phase shift proportional to frequency (without contributing to the scanning-loss), and two additional envelope frequencies that determine those details in the scanning effects that arise from the relative importance of pure memory and pure anticipatory components of scanning.

Some results of our curve-fitting game can be seen in Fig. 11; the calculated curve here was calculated from a simple three parameter scanning function by arbitrarily (1) matching the location of the first dip in Woodward's data by setting $w_c = 5.6 \mu\text{m}$, (2) setting the asymmetry parameter $a = 0.1$, and (3) using the simplest

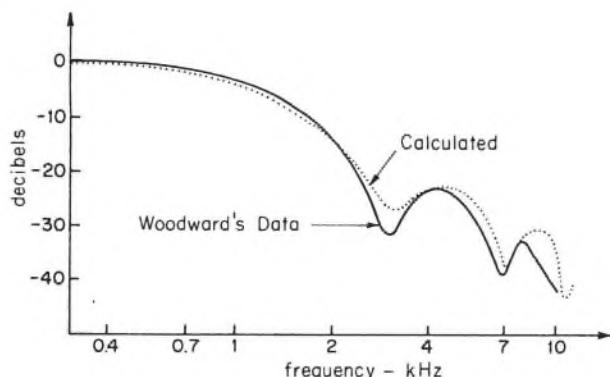


Fig. 11. Comparison of experimental and calculated scanning loss.

possible envelope correction function so that $|\bar{E}(f)|^2 = 1 + (f/f_e)^2$, in which the envelope frequency $f_e = 850$ Hz. A better match could be obtained by changing some of these arbitrary choices.

The weighting function corresponding to $f_e = 850$ Hz, $w_c = 5.6 \mu\text{m}$, and $a = 0$ is shown in Fig. 10, where the envelope function was chosen to model pure anticipatory scanning. The greatest weight is placed on the modulation located at the leading edge of the contact region. The weighting steadily decreases as one moves to the left toward the trailing edge. The weight placed on the modulation in the pileup region is relatively large near the leading edge but falls off exponentially as one moves away from the contact region to the right. Thus, according to the theory used here, pileup of plastic significantly increases the size of the effective aperture through which the stylus scans the modulation. There is an important distinction in this theory, however, between the contact region and the pileup region. The theory indicates that the critical frequencies of the dips are determined by the average width w_c of the contact region proper, independently of what is happening in the pileup region. In contrast to this, the frequency at which envelope loss starts to contribute significantly to the total scanning loss depends only on the decay constant $v/(2\pi f_e)$ of the exponential decay in the pileup region.

CONCLUDING REMARKS

The results described in this paper can be divided into two groups. In the first are those results that are valid regardless of the physical details of the stylus-groove interaction. One example of these is the fact that the frequency-response function can always be factored uniquely into a pickup function and a scanning function. Another example is the fact that the scanning function is always equal to the Fourier transform of the weighting function that we use to calculate the stylus displacement due to groove modulation when the pickup loss can be neglected.

The second group of results, in contrast, contains conclusions that were not deduced, but simply guessed. An example of these is the scanning function that we used to fit Woodward's data. There exist an infinite number of other scanning functions that would also match Woodward's data. This is true because (1) the data cover only a finite frequency range, and (2) the data give only the modulus of the scanning function. Phase information is missing. In fact there do not appear to be any data in

the literature giving the phase shift that accompanies pure scanning loss. Since phase data are absent and the frequency range is finite, the Fourier transform cannot be used by itself to calculate the weighting function from scanning-loss data. Some arbitrary assumptions must be used to choose a particular phase shift characteristic and a particular extrapolation of the data outside their frequency range to get unique results. The arbitrary assumptions that we used here are implicit in the special way we successively built up a plausible weighting function to match Woodward's data. Some other method might be found that is more plausible and direct and that yields different conclusions about the weighting function. More experimental results are clearly needed to find out if the weighting function used in the present theory is trustworthy.

In any case, a few final words are in order about what the pickup designer can do to control scanning loss. There are only two pickup design variables that influence scanning loss: tracking force and stylus geometry. Experimental results and the theory presented here both indicate that the critical scanning-loss frequency is raised either by reducing the tracking force or by decreasing the effective tip radius of the stylus. According to the above theory this raising of the critical frequency occurs simply because the width of the stylus-groove contact region is decreased. At present the only theory available for predicting this contact width is Hertz's theory [4], [5], [6] of static contact between two frictionless linearly elastic bodies. Needless to say, it is amazing that Hertz's theory has given as good of results as it has, since the stylus-groove contact is dynamic, involves friction, and yields groove stiffnesses that vary with tracking force in a non-Hertzian way [15].

ACKNOWLEDGMENT

Professor F. V. Hunt contributed encouragement and thoughtful criticism to this research, which was supported in part by the Office of Naval Research.

REFERENCES

- [1] J. M. Eargle, "Performance Characteristics of the

Commercial Stereo Disc," *J. Audio Eng. Soc.*, vol. 17, pp. 416-422 (August 1969).

[2] J. R. Sank, "Bi-Radial and Spherical Stylus Performance in a Broadcast Disc Reproducer," *J. Audio Eng. Soc.*, vol. 18, pp. 402-406 (August 1970).

[3] C. Kornei, "On the Playback Loss in the Reproduction of Phonograph Records," *J. Soc. Motion Pict. Engrg.*, vol. 37, p. 569 (1941).

[4] H. Hertz, *Miscellaneous Papers* (MacMillan and Co., New York, 1896), Chapters 5 and 6.

[5] S. Timoshenko and J. N. Goodier, *Theory of Elasticity*, (McGraw-Hill Book Co., New York, 2nd ed. 1951, pp. 372-377).

[6] D. H. Cooper, "Tables of Hertzian Contact-Stress Coefficients," Report R-387, Coordinated Science Laboratory, University of Illinois (1968).

[7] F. G. Miller, "Stylus-Groove Relations in Phonograph Records," Doctoral dissertation, Harvard University (1950). Issued as Tech. Memo No. 20, Acoustics Research Laboratory, Harvard University (1950). Only photocopies are now available, as Ref. No. PB125454, from Office of Technical Services, U.S. Dept. of Commerce, Washington, D.C. A brief summary of the results was presented in F. V. Hunt, *Acustica*, vol. 4, p. 33 (1954).

[8] P. Kantrowitz, "High-Frequency Stylus-Groove Relationships in Phonograph Cartridge Transducers," *J. Audio Eng. Soc.*, vol. 11, p. 250 (July 1963).

[9] C. R. Bastiaans, "Factors Affecting the Stylus/Groove Relationship in Phonograph Playback Systems," *J. Audio Eng. Soc.*, vol. 15, p. 389 (1967). See also Comments by D. A. Barlow, *J. Audio Eng. Soc.*, vol. 16, p. 364 (1968).

[10] J. V. White, "A Linear Theory of Phonograph Playback," *J. Audio Eng. Soc.*, vol. 19, pp. 94-100 (February 1971).

[11] M. S. Lighthill, *Fourier Analysis and Generalized Functions*, (Cambridge University Press, New York, 1958).

[12] J. V. White, "Mechanical Playback Losses and the Design of Wideband Phonograph Pickups," *J. Audio Eng. Soc.*, vol. 20, pp. 265-270 (May 1972).

[13] R. Bracewell, *The Fourier Transform and Its Applications*, (McGraw-Hill Book Co., New York, 1965), p. 366.

[14] G. Dickopp, H. Klemp, H. Redlich, and E. Schueler, "A Mechanical Disc Recording and Reproducing System with High Storage Density and High Rate of Information Transmission," *J. Audio Eng. Soc.*, vol. 18, p. 618 (December 1970).

[15] J. V. White, "An Experimental Study of Groove Deformation in Phonograph Records," *J. Audio Eng. Soc.*, vol. 18, pp. 497-506 (October 1970).

THE AUTHOR

James V. White was born in Hammond, Indiana in 1941. He received a BS in electrical engineering from Northwestern University in 1964, an SM in engineering in 1965 from Harvard University and a Ph.D. in engineering in 1970, also from Harvard, for which his major field was acoustics and signal processing.

While employed by the Jensen Mfg. Co. from 1961 to 1964 through Northwestern University's industrial

cooperative plan, he designed transducers for earphones and horn-loaded loudspeakers, and developed methods for testing the fidelity of sound reproduction in rooms. During the summer of 1970 Dr. White was a post-doctoral fellow at Harvard in the Division of Engineering and Applied Physics. He is now an Assistant Professor of Mechanical Engineering at Stevens Institute of Technology.

Tracing Distortion Correction*

E. G. TRENDELL

EMI Central Research Laboratories, Hayes, Middlesex, England

An analog shift register as a controlled-variable delay provides an accurate simulation of the tracing distortion normally present in phonograph records. Used as correction prior to lacquer cutting, an audibly effective improvement is possible. The determination of the control characteristics required to correct over a large frequency range is described.

INTRODUCTION: In the process of lacquer cutting and the subsequent replay of the record a discrepancy exists between the cutter movement and that transcribed by the pickup stylus. This tracing distortion can be partly attributed to the inability of the spherical (or elliptical) replay stylus to follow a groove cut by a chisel-shaped tool. Analysis of simple geometric distortion, which has been extensively dealt with elsewhere [1]–[9], follows two main approaches, 1) that of Fourier components and 2) as an advance and delay error with respect to the recorded signal. Experimental evidence indicates that other factors are involved, particularly tracking error and the yield of the material. However, it can be assumed that the overall effect is close to delay modulation and, as such, causes intermodulation between any two signals present. It follows that by the introduction of an exactly equal but opposing distortion before cutting it becomes possible to minimize the overall modulation. The most satisfactory method of producing an exact inverse distortion is by means of a variable delay line. Previous corrective systems have incorporated electromagnetic delay lines with taps at discrete positions. Now that an analog shift register is readily available, it is possible to obtain an infinitely variable controlled delay. The time differences between pickup reading point and the cutting edge can thus be

simulated. To investigate its application and advantages, such a system was set up providing an accurate simulation and thus correction for tracing distortion.

Consider the interaction produced when the input signal consists of two distinct bands, one with low- and the other with high-frequency content. As the lower frequency signal produces the larger effective delay, it can be said to cause most of the intermodulation distortion of the higher frequency signal. To provide correction to the higher frequency signal, it is passed through the variable delay which, in turn, needs to be controlled from some function of the lower frequency signal.

Because of the previously mentioned unknown factors all correction, at least initially, needs to be adjusted empirically. It is necessary to determine the precise nature of the required controlling function, how it varies with disk diameter, frequency, amplitude, and relative phase or timing. Thus a weighting network, to be applied to the lower frequency signal, can be defined that will automatically provide the correct compensation for all inputs and conditions.

Initial experiments involved only two tones and examination of the ensuing intermodulation products. This artificial signal is ideal and necessary for adjustment, but in order to assess performance, a musical source is essential. Maximum intermodulation distortion is thought to occur when high-level low frequencies and high frequencies are present at the same time. It becomes apparent

* Presented March 3, 1976, at the 53rd Convention of the Audio Engineering Society, Zurich, Switzerland.

as unwanted sidebands of the high-frequency signal. To obtain a meaningful audible effect, music modulated in a single sideband suppressed carrier mode is used as the high-frequency signal together with similar, not identical, music for the low-frequency or baseband source. This effectively separates the "interfering" from the "interfered with" signals. After demodulation the audible distortion products can be determined. Thus applied to each input of the lacquer cutting amplifier there is a baseband plus a single sideband signal. A common carrier frequency is utilized for both left and right channels. This arrangement is shown in schematic form in Fig. 1, and the resulting spectrum in Fig. 2 and 3.

If the baseband input signals are designated f_1 and f_2

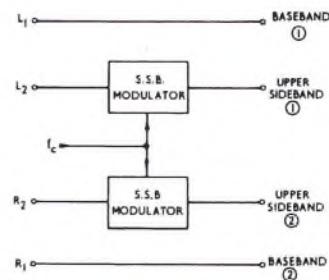


Fig. 1. Derivation of test signal.

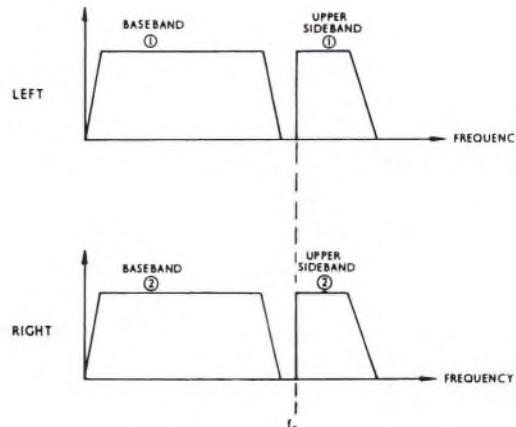


Fig. 2. Resulting spectrum.

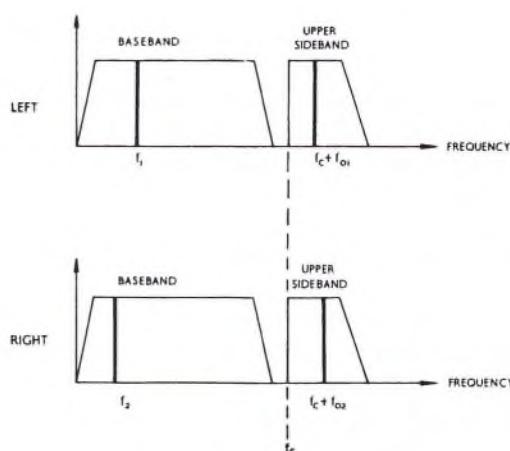


Fig. 3. Spectrum with specific components.

with $f_c + f_{01}$ and $f_c + f_{02}$ as the upper sidebands in the left and right channels, the components present in a delay or angle modulation that are of most interest are

$$[(f_c + f_{01}) \pm f_1], \quad [(f_c + f_{01}) \pm 2f_1]$$

1st order

2nd order

$$[(f_c + f_{01}) \pm f_2].$$

cross modulation

After single sideband demodulation these appear as

$$[f_{01} + f_1], \quad [f_{01} + 2f_1], \quad [f_{01} \pm f_2].$$

Fig. 4-7 are actual spectrum plots that occurred for an input signal consisting of two tones, one at 1 kHz, and the other at 17 kHz. Fig. 4 shows the two fundamentals together with sidebands that are produced by the cutting and replay process. Fig. 5 shows the output of a frequency discriminator when the above signal is applied to it. Figs. 6 and 7 are like Fig. 5 but with the first- and second-order tracing corrections, respectively.

THE DELAY MODULATOR

An analog shift register (bucket brigade) is used as a controlled-variable delay or delay modulator. Here the input signal is sampled with each sample being successively transferred along a line of buffered capacitors on a clock command. Time delay is determined by the clock frequency which, in turn, is made dependent on a control voltage.

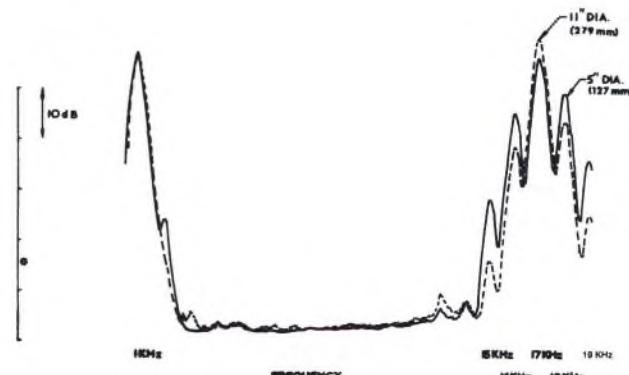


Fig. 4. Spectrum plot of pick-up output.

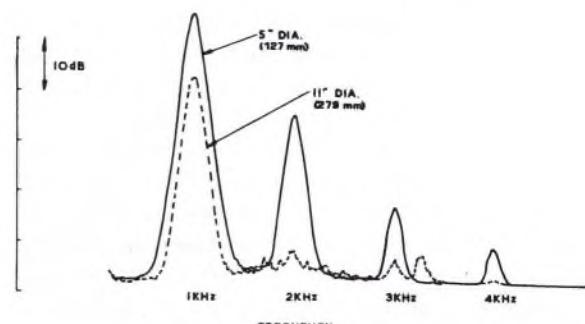


Fig. 5. Output of frequency discriminator.

The basic law of an analog shift register is

$$\text{delay } \tau = N/(2f)$$

where N is the number of "buckets" and f the clock frequency.

Thus the delay is proportional to the reciprocal of the clock frequency, which is derived from a voltage-controlled oscillator. This has a linear input voltage/frequency characteristic.

If V is the input to the voltage-controlled oscillator, then

$$\tau \propto 1/V.$$

For linear control within the required operating range of $\pm 50 \mu\text{s}$ the overall input is required to be proportional to $1/V$. This is achieved by means of an integrated variable conductance divider circuit. The block schematic is shown in Fig. 8. Following the variable-delay unit a low-pass filter removes the clock frequency components.

METHOD OF MEASUREMENT

For a test signal two separate tones were combined and applied to the lacquer cutting amplifier via the delay modulator as shown in Fig. 8. One tone, a high frequency, was maintained constant, while the other, a lower frequency, was varied over a range of frequencies and levels. It was intended that any distortion of the high-frequency signal that occurred in the cutting/replay process could be taken as a measure of the intermodulation. A pickup, which replayed the lacquer about half a revolution away from the cutter, provided a performance monitor. Output from the pickup, through a band-pass filter passing only the high-frequency signal plus sidebands, was applied to a phase-lock loop demodulator. Its output then indicated the extent and order of the intermodulation distortion. The

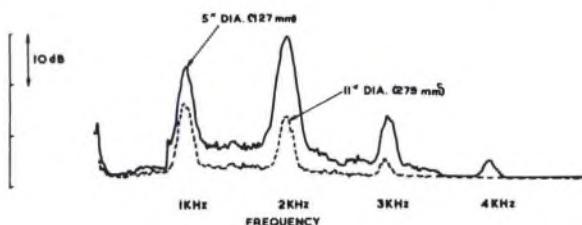


Fig. 6. Output of frequency discriminator, with 1st order correction.

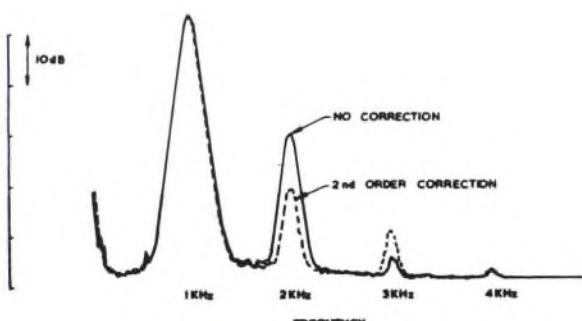


Fig. 7. Output of frequency discriminator with 2nd order correction.

control for the delay modulator was derived from the lower frequency input, with facilities to alter both amplitude and relative phase angle. In this manner adjustment was possible while monitoring the results from the pickup output. For each input frequency and at each disk diameter the first-order intermodulation component was reduced to a minimum and note taken of the required compensation amplitude and phase. By these means the complete amplitude/frequency and phase/frequency networks were defined which when implemented would be satisfactory for all input signals. To increase the discrimination between the intermodulation products and the background noise, the phase-lock loop output was applied to a frequency selective circuit tuned to a particular component as shown in Fig. 9.

For most of the record cut only the first-order intermodulation was of significance, but toward the center of the disk the second order became important, particularly when the first-order component had been reduced by compensation. In order to correct for second-order effects, a squaring circuit was introduced into the control and its appropriate amplitude/frequency and phase/frequency weighting network determined. Also there exists cross-modulation between left and right channels, and this again produces unwanted sidebands. Correction can be effected by some cross coupling in the left and right controls.

For all these compensation networks the amplitude/frequency and phase/frequency characteristics required empirical determination. The complete schematic is shown in Fig. 10 with both first and second order correction and cross coupling.

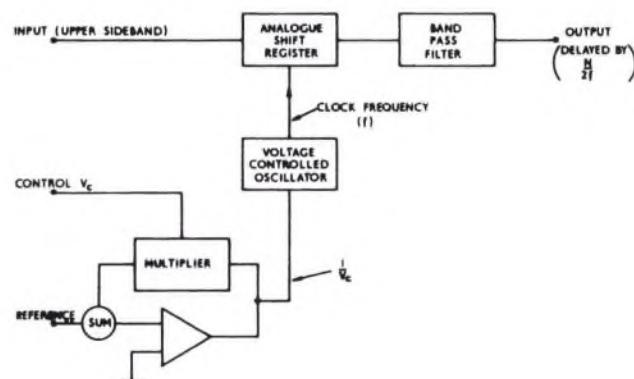


Fig. 8. Delay modulator.

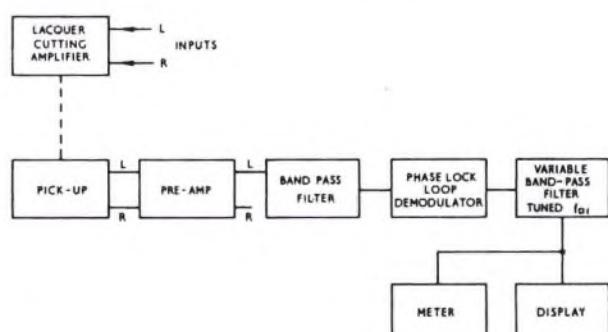


Fig. 9. Measurement of intermodulation.

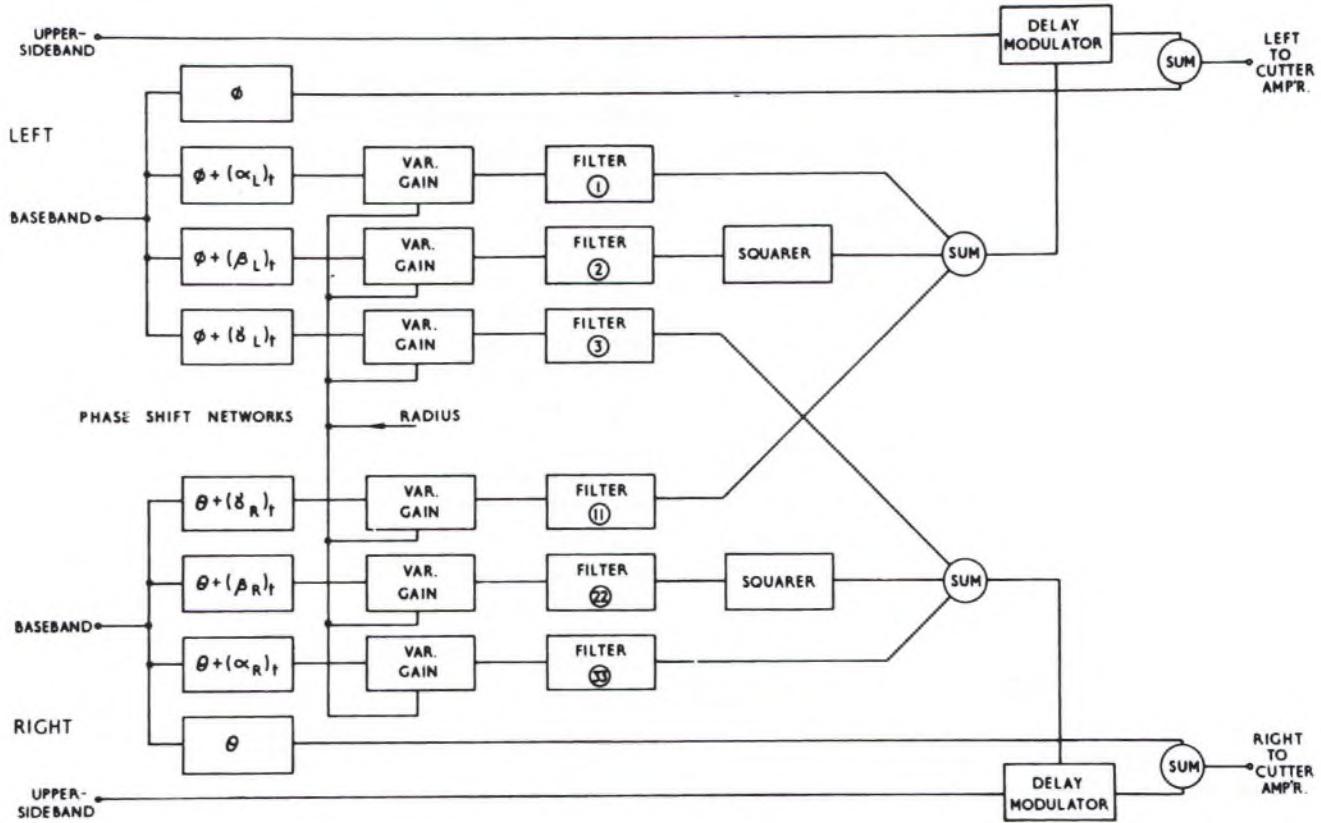


Fig. 10. Complete schematic of correction circuit.

RESULTS AND CONCLUSIONS

For complete compensation of the intermodulation distortion that occurs in the lacquer cutting/record replay process the procedure can be summarized as follows.

- 1) The necessary control amplitude is directly proportional to the baseband level.
- 2) The control amplitude varies with the diameter of cut, as shown in Fig. 11.
- 3) The control frequency response is very close to the standard RIAA replay characteristic. However, the relative phase response needs some modification.
- 4) It is possible to reduce the first-order component by between 20 and 25 dB for baseband frequencies in the range 100 Hz to 2 kHz.
- 5) From the outside down to about 7 inches (178 mm) in diameter only the first-order intermodulation components are of importance, below this the second order become appreciable. A reduction of 10–15 dB is possible by means of the second-order compensation.
- 6) The relative level of the left-right crossmodulation components is about 14 dB less than the intermodulation.
- 7) There appear to be only small differences between the modulation sidebands produced when either the lacquer or a vinyl pressing made from the lacquer is played.
- 8) All correction is optimized for one particular shape and type of pickup stylus. Any variation in size and shape will reduce the overall improvement. The extent of this needs further investigation.
- 9) An audible improvement can be heard when listening to the demodulated upper sideband for a music

modulated single sideband test signal if the simulation correction is in circuit.

ACKNOWLEDGMENT

The author wishes to thank his colleagues Cyril Foxcroft for many helpful discussions and Malcolm Cochran who did most of the work.

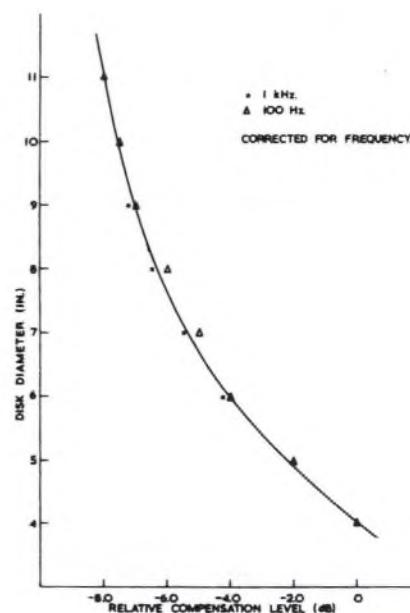


Fig. 11. Variation of compensation with disk diameter.

REFERENCES

- [1] W. D. Lewis and F. V. Hunt, "A Theory of Tracing Distortion in Sound Reproduction from Phonograph Records," *J. Acoust. Soc. Am.*, vol. 12, p. 348 (1941).
- [2] M. S. Corrington, "Tracing Distortion in Phonograph Records," *RCA Rev.*, p. 241 (June 1949).
- [3] J. B. S. M. Kerstens, "Mechanical Phenomena in Gramophone Pick-ups at High Audio Frequencies," *Philips Tech. Rev.*, vol. 8, 3 (1956/57).
- [4] J. Walton, "Wireless World," p. 353 (July 1961).
- [5] E. C. Fox and J. G. Woodward, "Tracing Distortion—Its Causes and Correction in Stereodisk Recording Systems," *J. Audio Eng. Soc.*, vol. 11, p. 294 (Oct. 1963).
- [6] 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, p. 111 (Apr. 1965).
- [7] D. H. Cooper, "Interaction of Tracing and Tracking Error," *J. Audio Eng. Soc.*, vol. 13, p. 138 (Apr. 1965).
- [8] S. Washizawa, T. Nakatani, and T. Shiga, "Development of Skew-Sampling Compensator for Tracing Error," *J. Audio Eng. Soc.*, vol. 21, pp. 630–634 (Oct. 1973).
- [9] C. L. Foxcroft, British Patents 1,089,291; 1,193,073; 1,118,381.

THE AUTHOR

Edward G. (Ted) Trendell was born in 1928 at Beaconsfield in Buckinghamshire, England. He obtained a B.Sc. (Eng.) degree in 1954 from the University of London.

Since then he has held the position of senior research engineer, with particular emphasis on audio techniques, at the Central Research Laboratories of E.M.I. Ltd., Hayes, Middlesex. His main interests have been in methods of automatic level control of audio signals in the form of limiting, compression, expanding, and compansion. His paper on the measurement of tape modulation noise, published in the journal (Vol. 17, No. 6, Dec. 1969.) has

been used to provide the basis of an I.E.C. standard. He has read a number of papers on quadraphonic systems both to Conventions of the Audio Engineering Society and to meetings arranged by other bodies. Recently he has been concerned with an examination of the finer details and inadequacies of lacquer disk cutting and control. At the 50th A. E. S. Convention in London he was a member of the organizing committee, chairman of two sessions and the visits subcommittee.

Mr. Trendell is a Chartered Engineer, a member of the A. E. S. and a member of the Institution of Electrical Engineers.

Groove Deformation and Distortion in Records*

D. A. BARLOW

Bingley, Yorkshire, U.K.

AND

G. R. GARSIDE

Department of Computer Science, University of Bradford, Yorkshire, U.K.

The elastic and plastic deformation of vinyl record compound under indenters of various profiles has been studied in large-scale tests. Curves of total depth of penetration versus load have been used to calculate the net distortion on record playback. In general, in the lower treble, net distortion is less, and in the upper treble, net distortion is greater than tracing distortion alone. This is important for attempts to reduce tracing distortion by inverse predistortion of the recorded signal. Nylon was also studied as a material with contrasting mechanical properties to vinyl. Further light has been shed on the nature of translation loss.

INTRODUCTION: In the playing back of records, tracing distortion is serious, as it can reach very high values. The original groove in the lacquer master is cut by a chisel-shaped cutter of 90° included angle, but the record is played back with a curved stylus. Thus it follows a different path from the cutter, and the difference, and hence the distortion, is most severe at high frequencies and at the lowest record speeds, that is, near the center of the record. By reducing the stylus radius, tracing distortion is reduced, and to avoid excessive deformation of the record groove, the lateral radius is increased, giving an elliptical or biradial stylus. Mathematical analyses of tracing geometry (assuming that the groove is rigid) have been made by Hunt [1], [2], Corrington [3], [4], and Cooper [5].

It is also known that there is considerable elastic and plastic deformation of the record groove under the load of the stylus [6]–[9], but the effect of this on the signal is not accurately known. Obviously it constitutes distortion which may add to or counteract tracing distortion. For static indentations, the elastic range is covered by the

Hertz equations, which show that the elastic limit is reached at a load of 3 milligrams on a 12.7- μm (0.0005-inch) radius indenter on a flat surface of vinyl record compound. No solution exists for the plastic range, especially for sliding indentation with friction. Any solution is likely to be very long and cumbersome to use.

Considerable information is available on the widths of tracks made by indenters, but little on depths, these being much more difficult to measure, especially on the very small scale of a record groove. Walkling [10] has measured the mechanical impedance of vinyl surfaces by means of a vibrating pickup, but there is no simple way of converting his results to the required curve for load versus depth of penetration under load. White [11] has devised a mathematical conversion which suggests that at very low loads, the groove is more compliant than given by the Hertz equations; at high loads, in the fully plastic range, his results suggest that the material is even less compliant than a fully elastic material would be. The apparent behavior at low loads might be due to flattening of the surface asperities prior to “bedding down” on to solid material; at high loads, residual internal stresses, caused by rapid cooling, may increase the apparent hardness.

* Manuscript received December 1977; revised May 1978.

The total depths under load, measured from the original surface, were therefore determined in large-scale tests over the appropriate range of loads, thus avoiding "noise" due to surface roughness of the groove. The material was slowly cooled to avoid residual stresses. Results represent the basic mechanical properties of the material without extraneous factors and enable the deformation and distortion to be calculated for record grooves. The deformation of the groove under load itself alters the load on the stylus, which in turn alters the deformation, etc., etc. The calculation is thus iterative, and a computer was used for this purpose.

1. MECHANICAL PROPERTIES OF RECORD COMPOUND

A preliminary investigation was made of the mechanical properties of the vinyl chloride-acetate co-polymer used for records. This contains about 16% of acetate to lower the processing temperature, together with small quantities of carbon black pigment, stabilizers, and lubricants (to assist removal from the mould), but no plasticizer. The material was in the form of blocks about 1 cm thick, made by compression moulding. All the tests were carried out at 18–20°C to avoid variations in properties due to temperature.

Compression tests were made on cylinders machined from the blocks, and curves are given in Fig. 1. The stress is plotted as nominal stress (= load/original cross-sectional area) and true stress (= load/current cross-sectional area). The latter gives a better indication of what the material is doing. The initial portion of the curve is elastic, the slope giving Young's modulus of elasticity. After yielding, the load which the specimen can support drops, and even after considerable deformation, (measured as contraction in height = (original height – final height)/original height × 100%), the original yield stress is not regained, that is, the material work-softens before eventually beginning to work-harden. This is unusual, although there are many materials which show very little work-hardening. This must affect the record distortion characteristics, and nylon was chosen as a contrasting material, as it is known to work-harden considerably. In this case, compression moulding cannot be used due to excessive oxidation, and injection moulding gives unsoundness in thick sections. The blocks were therefore cast from hot monomer, which polymerized in the mould to nylon 66. Similar compression tests were made and

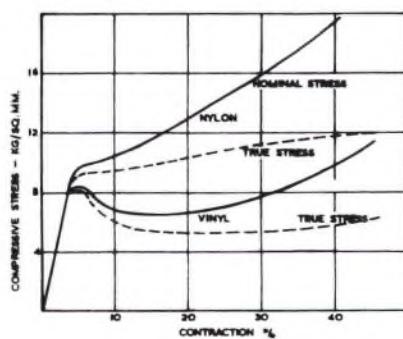


Fig. 1. Compressive stress-strain curves on vinyl and nylon.

curves are given in Fig. 1, and show typical work-hardening from the onset of yielding.

A large number of static and sliding indentations were made on the two materials, which showed that they were fully elastic at sufficiently low loads and did not behave differently in principle from other materials. This has been fully reported elsewhere [12].

2. DYNAMIC SLIDING INDENTATION TESTS

Blocks with a flat as-moulded surface were mounted on the faceplate of a lathe. The surface speed was 35–45 cm/s, the properties varying little over the range of interest. The indenter of hardened steel was mounted on a rigid lever arm pivoted on the saddle. The load was applied by means of pulley, cord, and weights. A dial gauge reading to 1/10 000 inch was mounted adjacent to the indenter, bearing on untested material adjacent to the tracks made by the indenter. The spherical indenter was of 1.27-mm radius, corresponding to a 12.7-μm (0.0005-inch) radius stylus. Indenters corresponding to biradial styli of 18/9 μm (0.0007/0.00035 inch) and 20/5 μm (0.0008/0.0002 inch) radius styli were also used. So-called elliptical styli are usually made by grinding flats on a conical stylus. The end is then polished to remove sharp corners and to produce the desired radii. The exact profile at the point of contact with the record may be uncertain. In the present case, the indenters were truly biradial and to the correct radii. An indenter equivalent to 70/7 μm was also included, representing the Shibata and similar styli.

Curves are given in Figs. 2 and 3 for 1, 2, 5, and 10 playings. Over most of the range of interest, the curves are straight on a log-log scale and the corresponding equations are given in Table I. The slope of the vinyl curves decreases with repeated playing, whereas for nylon the slope does not change. This is a function of the greater work-hardening of the nylon, for which the slope is conveniently 1. The 70/7-μm equivalent indenter gave results very similar to the 18/9-μm indenter.

As a check on these large-scale tests, tracks were made on the central flat portion of a record at loads of 4, 6, 8, 10, and 12 grams, using a 25.4-μm radius stylus at a record speed of 36 cm/s. The widths of the tracks were measured, using a microscope with graticule eyepiece and

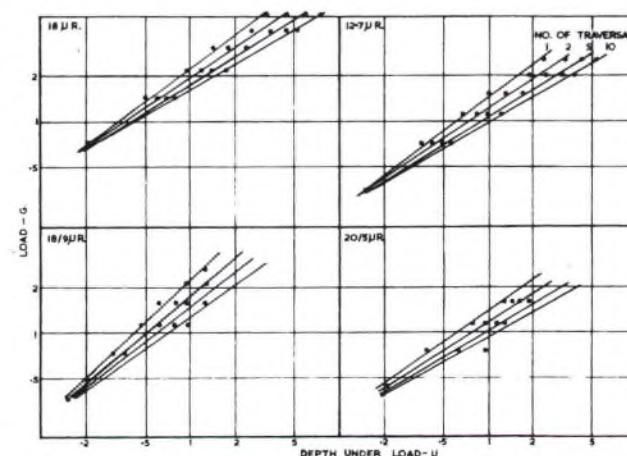


Fig. 2. Dynamic load versus penetration, vinyl.

vertical illuminator. These track widths are compared with the equivalent widths from the large-scale tests, (Table II). Agreement is good, considering the difficulty of making accurate measurements on very small tracks.

Lewis [13] has produced excellent micrographs of record grooves, using the scanning electron microscope. These provide a further check on the present tests. A playing weight of 5 grams was used, corresponding to 3.54 grams on each groove wall. The corresponding track width expected from large-scale tests for one playing would be 14.8 μm . Lewis's maximum and minimum track widths for one playing of a groove with a sine wave test signal are 17.7 and 11.8 μm , giving a mean of 14.8 μm .

3. DEFORMATION OF CURVED SURFACES

To measure depths of penetration at high speed on curved surfaces is difficult, so slow-speed tests were made to obtain ratios of (depth on curved surface) to (depth on flat surface). Specimens were moulded with convex and concave profiles of various radii, and the depths were measured while traversing under load, using lever arm and dial gauge as before. The largest possible indenter was used to give the greatest accuracy. The spherical indenter was of 2.38-mm radius, representing a 12.7- μm radius stylus, and corresponding biradial indenters were used. There was considerable scatter with the vinyl. To reduce this, results were repeated several times, and between each

set of 10 traverses, the depth on the adjacent flat surface was measured. This also reduced possible scatter due to temperature and humidity variations and variation in material properties from point to point. The scatter on the vinyl, in contrast to the nylon, was doubtless due to the lack of work-hardening. This took the form of instability or juddering with increasing ellipticity of the indenter. This gave some variation in track depth. With excessive loads, the biradial indenters dig in and act as a cutting tool, with the formation of a chip.

The results represent over 20 000 traversals. Values may be plotted as the ratio (depth on curved surface) to (depth on flat surface) versus load and as ratio versus radius of profile. These form a cross check, as the ratio versus load curves must be drawn through scattered points such that the ratio versus radius curves are smooth. The latter curves are given in Figs. 4–11. In general, the ratios are lower for the nylon, that is, geometry has less effect on depth of penetration. This again is doubtless due to the work-hardening of the nylon.

4. COEFFICIENT OF FRICTION

The coefficient of friction varies according to measuring conditions and may vary considerably over different parts of a given record. Rangabe [14] has measured friction in

Table II. Comparison of tracks from pickup with large-scale tests.

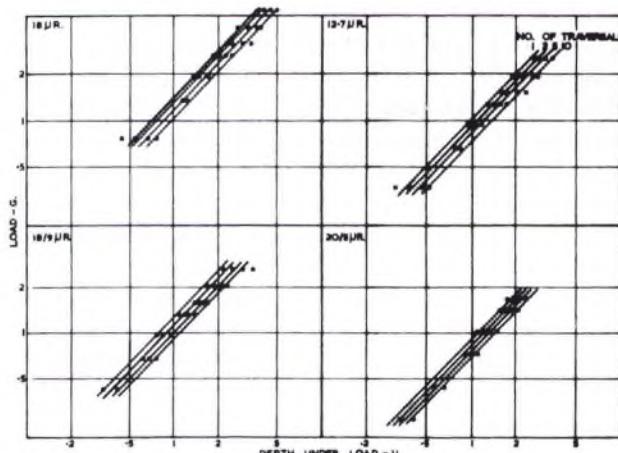


Fig. 3. Dynamic load versus penetration, vinyl.

Table I. Equations for dynamic depth versus load.

	Depth [μm] = a (Weight w [gram]) ^b			
	18- μm Radius	12.7- μm Radius	18/9- μm Radius	20/5- μm Radius
Vinyl				
1 playing	$0.30w^{1.46}$	$0.59w^{1.46}$	$0.43w^{1.03}$	$0.74w^{1.25}$
2 playings	$0.33w^{1.57}$	$0.70w^{1.57}$	$0.50w^{1.12}$	$0.91w^{1.36}$
5 playings	$0.38w^{1.68}$	$0.85w^{1.68}$	$0.57w^{1.22}$	$1.09w^{1.47}$
10 playings	$0.42w^{1.79}$	$1.00w^{1.79}$	$0.64w^{1.33}$	$1.29w^{1.58}$
Nylon				
1 playing	$0.66w$	$0.94w$	$0.83w$	$1.13w$
2 playings	$0.75w$	$1.06w$	$0.93w$	$1.25w$
5 playings	$0.84w$	$1.18w$	$1.03w$	$1.38w$
10 playings	$0.92w$	$1.30w$	$1.13w$	$1.50w$

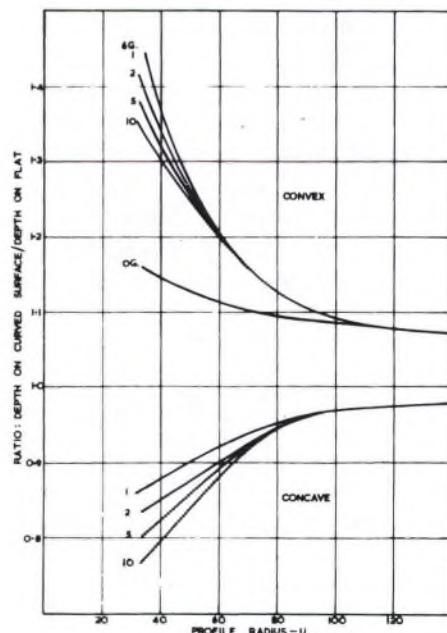


Fig. 4. Depth ratio versus profile radius, vinyl 18- μm radius stylus.