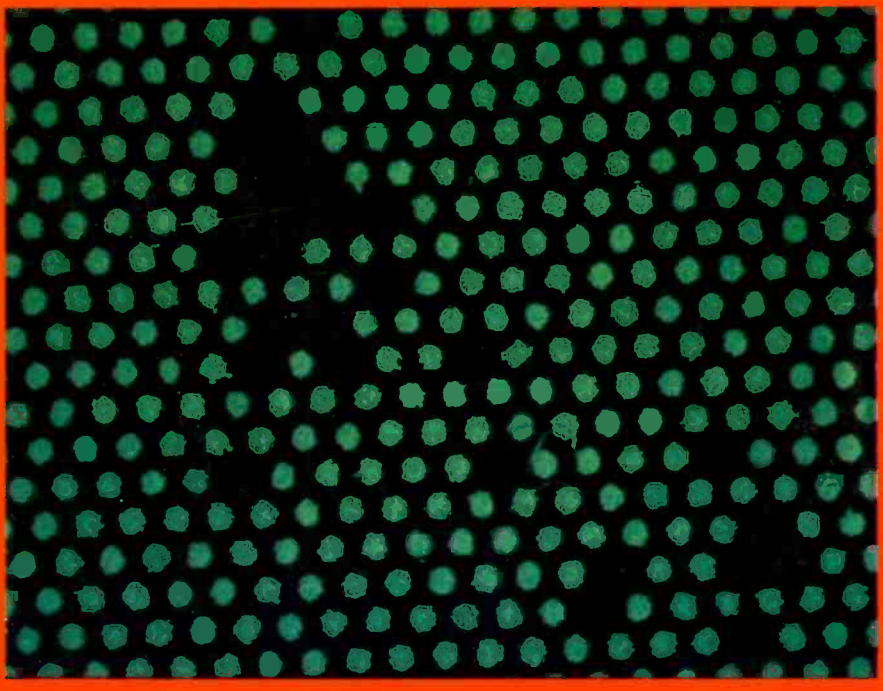
RCA Review



Measuring Phosphor Screen Adherence

September 1984 Volume 45 No. 3

RCA RCI 45(3) 335–552 (1984)

Our cover photo shows a portion of a phosphor screen (magnified 25×) deposited on a 3 × 3 inch piece of faceplate glass after testing to measure adherence. Only the green phosphor dots (shown here illuminated) have been deposited. The sample shown was used in an investigation of the factors affecting the adherence of phosphor screens to faceplate glass in the manufacture of TV tubes, as described in the paper by Larach and McGowan in this issue of *RCA Review*.

RCA Review, published quarterly in March, June, September, and December by RCA Laboratories Princeton, New Jersey 08540. Entered as second class matter July 3, 1950 under the Act of March 3, 1879. Second-class postage paid at Princeton, New Jersey, and at additional mailing offices. Effective January 1, 1983, subscription rates as follows: United States. one year $12.00, two years $21.00, three years $27.00; in other countries: one year $14.00, two years $24.50, three years $31.50. Single copies up to five years old $5.00

RCA Review

RCA Review (ISSN 0033-6831) is a technical journal published quarterly by RCA Laboratories in cooperation with the subsidiaries and divisions of RCA.

Contents

September 1984 Volume 45 Number 3 ISSN 0033-6831

Adherence of Phosphor Screens

S. Larach and J. E. McGowan
RCA Laboratories, Princeton, NJ 08540

Abstract—Properties of poly(vinyl alcohol), faceplate glass, phosphor slurries, and their interfaces have been investigated with respect to wet adherence of phosphor screens. Through measurements of bulk and interfacial properties, correlations of wet adherence were found with hydrodynamic parameters of PVA, using a new method of measuring adherence. The reasons for etching and precoating are discussed, and evidence is presented for a model involving glass surface silanol linkages to the PVA precoat, as well as the effects of the hitherto unsuspected glass-surface impurity layer due to etching. It has also been established for faceplate glass that each successive phosphor slurry sees a different surface in screening, the largest change being for the second phosphor deposited.

1. Introduction

While the PVA-dichromate system has been employed for many years in the world-wide production of color television picture tubes, there is a great paucity of data on the basic aspects of phosphor adherence to kinescope face-plates, a matter of great importance. The present study had three major aspects: (1) to determine the effect of various slurry additives on adherence, (2) to investigate basic screen adherence properties, in order to understand the processes that occur in screening and (3) to improve, if possible, the screen adherence of phosphors.

In defining the problem, we must also define the terms used. By *adherence,* we mean a phosphor screen, dot or line, or portion thereof, that is bonded to a glass substrate. This does not basically refer to particle-to-particle bonding, but rather the bonding of a particular geometric ensemble of phosphor particles to the glass

substrate and, in particular, to faceplate glass. In addition, we define *wet adherence* as the adherence of a phosphor screen after it has been wet with water, i.e., after development.

The experimental procedure for screening is shown in Fig. 1. The left portion of the figure involves the preparation of the substrate, faceplate glass (FPG); the right portion of the figure involves the chemical aspects of the phosphor slurry preparation/sensitization. The center portion of the figure treats of the screen application steps through developing, followed by wet adherence measurement by the jet impingement method. Although these steps are inter-related, different areas were studied by different techniques in order to arrive at a coherent model for phosphor screening.

Over the years, much practical lore has accumulated on phosphor screen adherence. Some beliefs are due to individual engineer's preferences or prejudices, but some constitute facts, more-or-less agreed to by the industry. These are usually based on invaluable experience in turning out a product, although the basic reasons for their use may have remained unknown. Two examples of these areas are etching of the substrate and use of a pre-coat prior to screening. These two particular aspects are part of the research reported here.

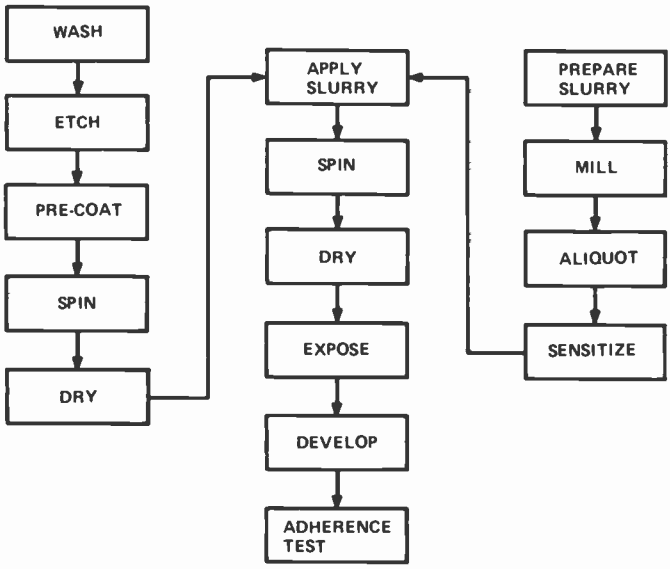


Fig. 1—Experimental procedure for screening.

In examining the overall problem of phosphor screen adherence, we decided to treat it as an interacting multi-interface system. As shown schematically in Fig. 2, we can consider a system comprising a section of faceplate glass (FPG) which forms interface I_1 and I_2, although some work was also done on the other interfaces. The way I_1 is affected by the etching treatment must also be considered in this analysis. The research involved the bulk properties of PVA, the FPG surface as affected by the etchant, the properties of dilute solutions of PVA, as used in precoats, and the surface chemistry of the glass. These will be treated in separate sections, as they relate to adherence.

The phosphor slurry was reduced eventually to its basic components: phosphor, PVA, water, and dichromate. This slurry is referred to as a model slurry in this paper. Phosphors were obtained from the RCA Video Components and Display Div., and the sensitizer was sodium dichromate. A variety of techniques were used in characterizing the various aspects of adherence.

This program was divided into several phases, the first of which was to evolve a method for studying adherence quantitatively. Adhesion measurements are best summarized by a statement from Mittal[1] in a 1981 paper: "Although there is available a cornucopia of adhesion measurement techniques, there is no single technique which can be recommended in all situations and which can be accepted by all those who have the need to measure adhesion. As a matter of fact, one should use that particular technique which best stimulates the usage conditions of the coating."

2. Adherence Measurements

To measure wet adherence of our screens, a zero-degree-spread jet

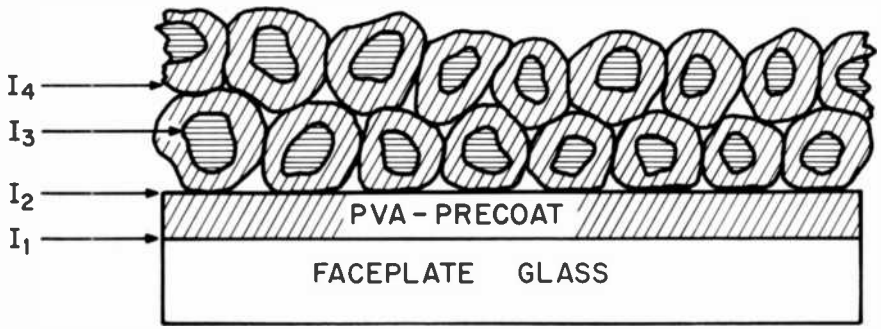


Fig. 2—Phosphor screen interfaces.

impingement method was evolved. In this method, a jet of de-ionized water, at a constant temperature, is impinged at constant pressure and time onto a wet screen that is a constant distance from the jet nozzle. The area of the screen lost is determined either by microscopic measurement or photometrically, and the wet adherence is taken as being inversely related to the area, i.e., the smaller the area of screen lost, the larger the adherence-factor. Relative adherence factor can then be plotted against some variable to compare the effects of changes or treatments.

An advantage of this method is that, with certain simplifying assumptions, values of "absolute" adherence can be calculated in ergs cm^{-2}. Both relative and absolute values of wet adherence are given in this paper. For a thorough treatment of the theory of particles in a water stream, the reader is referred to A. D. Zimon.[3]

Fig. 3 shows a typical plot for screen adherence, where the total jet kinetic energy is plotted against hole area for different times. The slope of the line yields a value of 8×10^7 ergs cm^{-2} for wet adherence for this particular case. Using a different and much more elaborate technique, Deryagin[2] has reported adherences of about 6

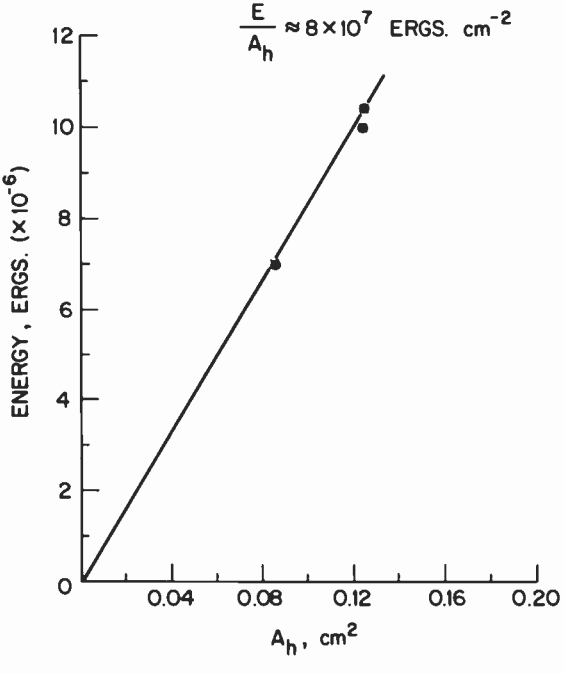


Fig. 3—Plot of kinetic energy vs hole area

$\times 10^5$ ergs cm^{-2} for polymer films on germanium surfaces, and about 3×10^4 ergs cm^{-2} for chlorinated PVC films on glass.

The mini-screening apparatus of Fonger[4] was adapted for use in our work (see Fig. 4). The holding tank for water was connected by ¼-inch stainless steel tubing to a pump capable of delivering water at 65 PSI at 21°C. The outlet side of the water pump was connected by ½-inch stainless steel tubing leading to a Jenkin's Bros. gate valve and an in-line valve reduction train. This consists of three Wilkerson brass reduction valves connected by ¼-inch stainless steel tubing to a Unijet zero-degree-spread nozzle, solid stream tip number 000019, 0.12 inch in diameter, obtainable from the Spraying Systems Co. Each Wilkerson reduction valve had a 0-60 PSI gauge attached to it.

To activate this system, the pump is turned on with a gauge pressure reading of zero. The Jenkins gate valve is opened, and by means of the speed control of the pump, the reading on the U.S. Gauge connected above the pump is raised slowly to 50 PSI. The first in-line Wilkerson reduction valve (nearest the pump) is set at 46; the next reduction valve is set at 40 PSI; and the last reduction valve before the solid stream nozzle is set at 36 PSI. The distance from the target is 7.8 cm. The target, a phosphor screen deposited on a 3 × 3 inch faceplate glass, is exposed and developed and set in an aluminum enclosure which is approximately a one foot cube, open at the top and at the jet-stream delivery side. The stream is permitted to hit the target for a given length of time. The target is dried under an IR lamp and the hole in the target is measured.

Nineteen-inch panels of virgin faceplate glass (FPG) cut into 3 × 3 inch sections were used in our experimental work. The glass sec-

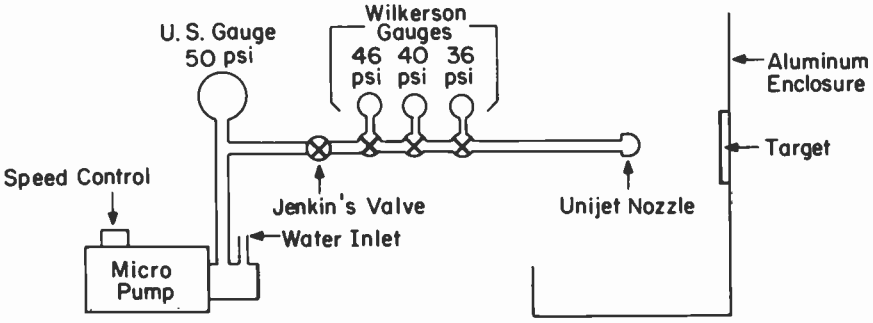


Fig. 4—Zero-degree-spread jet system for measuring adherence.

tions are first washed with a mild soap (Ivory) and warm tap water, followed by thorough rinsing in warm tap water and then in de-ionized water. The glass is then etched in a 5% solution of ammonium bifluoride for 30 seconds and rinsed with de-ionized water. While the glass is still wet, a precoat consisting of 0.5% poly(vinyl alcohol) at a pH of 2.7 is applied. This is done by using a Manostat syringe to apply the PVA uniformly over the glass surface. During PVA application, the glass rotates on a spinner at 7 rpm under a filtered 250 watt IR lamp 7 inches above the platform. As soon as the 2 ml of precoat PVA has been applied to the glass, the speed of rotation is increased to 96 rpm and the precoat is dried under the IR lamp for 5 minutes. Ten ml of phosphor slurry containing sodium dichromate is applied over the dry precoat on the faceplate glass at 7 rpm. The phosphor slurry consisted of the following formulation:

PVA 540/Total Solution = 3.5%
Phosphor/PVA = 12.5%
Pluronic/Phosphor = 0.1%
Phosphor/Total = 30%

After the phosphor slurry has been applied, the speed of rotation is increased to 96 rpm and the coating table holding the spinner is tilted to an angle of 105°. The IR lamp is attached to and tilts with the spinner, maintaining a constant normal incidence to the coating. The drying time is six minutes.

After the drying has ended, the spinner is tilted upright and the 3 × 3 inch glass section is removed. For the purpose of adherence testing, the entire phosphor coating is exposed using a bare 200 watt Hanovia X2-Hg compact-arc-lamp without a collimator. The source-to-coating distance is 13 inches, while the time of exposure can be varied.

After exposure, the coating is developed using 47°C deionized water for 30 seconds at 15 psi over a distance of 1 foot.

While still wet, the coating is set in the aluminum enclosure for adherence testing and is subjected to the jet stream. The coating is dried and the hole in the coating is measured for adherence calculation.

Fig. 5 shows wet adherence measurements for an RCA green-emitting phosphor, made into a model slurry, as a function of exposure time. We see, in this semi-log plot, the monotonic increase of adherence with exposure time until at about 55 seconds exposure, there is a sharp increase in adherence. Two reactions are proceeding; one at shorter exposure times with a specific rate constant of 4×10^{-4} sec^{-1} and a half-life of 1.7 msec, and one at larger

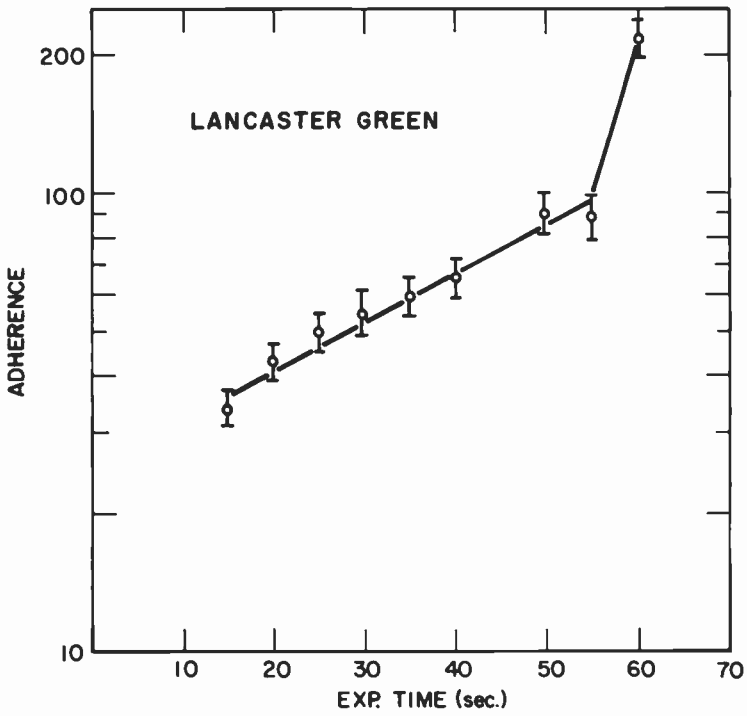


Fig. 5—Wet adherence vs exposure time.

exposure times with a specific rate constant of about 2.7×10^{-3} sec^{-1} and a half-life of about 25.6 msec.

3. Bulk PVA Properties

Poly(vinyl alcohol), PVA, is utilized in our manufacturing process both as a pre-coat and as a major part of the phosphor slurry, which is sensitized with dichromate. By poly(vinyl alcohol), we are referring to a material structured of linear molecules, A-$(CH_2OH)_p$-B, where p is the degree of polymerization and the end groups A and B are small and chemically inert. The average sterochemistry of the vinyl groups with respect to each other determine the *tacticity*. However, PVA, usually prepared from poly(vinyl acetate), may have minor structural complications because of side reactions. These are summarized in Table 1.[10] While PVA easily loses water when heated at 60°–100°C, this does not provide complete dehydration. Complete dehydration of PVA has been reported[11] by exposure for at least five hours to a suspension of calcium hydride in dry

Table 1—Main Structural Deviations from $-(CH_2CHOH)_p-$ Contained in Poly(vinyl alcohol) (From Pritchard[10])

Structure	Likely percent of monomer units	Cause
$-CH(OR)CH_2-$	0–5% (ethers)	Chemically incomplete reaction during transformation to PVA
$-CH(OCOR')CH_2-$	0–5% (ethers)	
$-CH_2CHOHCHOH-CH_2CH_2CHOH-$	0–2%, vinyl esters 0–1/2%, vinyl ethers	Occasional condensation of monomer by α-carbon
$-CH_2COH(CH_2CH_2-OH-)CH_2CHOH-$	0–0.05%	Elimination of hydrogen atoms at α- and β-carbon to start branches
$-CH(CH_2CHOH-)-CHOH-$	0–0.05%	
$-CH_2COH(CH_2CHOH\ CH_2CH_2OH)-$	Not easily estimated	Termination of chain end by abstraction of hydrogen on penultimate α-carbon
$-CH_2(C-O)-$	0–0.02%	Oxidation by air, especially catalyzed by bases
Cyclic ketal in single chains, and ketal between two chains	Very small	Condensation of oxo and hydroxy groups
Ether formation in single chains, and cyclic ether in single chains	Very small	Dehydration
$-CH=CH-$	High if polymer heated in right medium	Dehydration
$-CH_2OH$ end group	(Ideal)	Termination of propagating chain by abstraction of hydrogen atom
$-CH(OH)X$ end group	0–50%	Termination by X
$-CHO$ end group	0–50%	Elimination of catalyst residue as HX, from end group; especially HX = RCO_2H, ROH, or H_2O
Interchain cyclic acetal	Not easily estimated	Condensation of aldehyde group at end of chain with hydroxy groups

pyridine at 117°C. The resultant dehydrated product was a black polymer containing long chains of CH = CH groups. To examine some of the basic properties of PVA, particularly in dilute solutions (the PVA concentration in precoats is about 0.5% or less), measurements of viscosity were carried out.

3.1 Viscosity

If force per unit area, s, causes a layer of liquid at a distance x from a wall to move with a velocity, v, the viscosity, η, is defined as the

ratio between the shear stress s and the velocity gradient $\partial v/\partial x$ or rate of shear, γ, so that

$$S = \eta \frac{\partial v}{\partial x} = \eta\dot{\gamma}. \qquad [1]$$

A very convenient way of characterizing PVA materials is by their viscosities. In our work, we used an Ubbelohde viscometer, in a thermostatted water bath held at 30°C. The initial PVA solution was less than 1% and was diluted after each run to the desired concentration. Thermal equilibrium times were taken into account throughout the experiments, and triple-distilled water was used as a reference standard before each measurement.

Specific viscosities, η_{sp}, were obtained from flow-time measurements for water and for the PVA solutions. The intrinsic viscosity, $[\eta]$, is defined* as

$$[\eta] = \lim_{c \to 0} (\eta_{sp}/c), \qquad [2]$$

where c is the PVA concentration in grams per deciliter. $[\eta]$ was obtained by extrapolating the reduced viscosity, (η_{sp}/c) plotted against c, to zero concentration. Fig. 6 illustrates our results for three different PVA materials.

Intrinsic viscosities characterize polymers where the individual polymer molecules can be taken as being essentially free of the influence of neighboring molecules. Again, this is of particular interest for PVA precoats, since dilute solutions are involved. We can compare our intrinsic viscosities with the intrinsic-viscosity-molecular-weight relationships reported for aqueous solutions of PVA. Using the equation of Matsumoto and Ohyanagi[5],

$$[\eta]_{30°C} = 4.25 \times 10^{-4} \overline{M}_w^{0.64} \qquad [3]$$

we obtain $[\eta] = 1.7$ for PVA540, compared to our 1.08, which is in good agreement considering our nonfractioned PVA.

Elias[6], in a review of available data, has published the equation

$$[\eta]_{25°C} = 7.31 \times 10^{-4} \overline{M}_w^{0.616} \qquad [4]$$

found to apply over a range of PVA molecular weights from 10^4 to 5×10^5. Using the Elias equation, we obtain $[\eta]_{25°C} = 0.97$ for our PVA540 material, again as against our $[\eta] = 1.08$ value, at 30°C.

* Note that $[\eta]$ is given in deciliters per gram, and is sometimes referred to as the "limiting viscosity number" in the revised clasification system recommended by IUPAC in 1952.

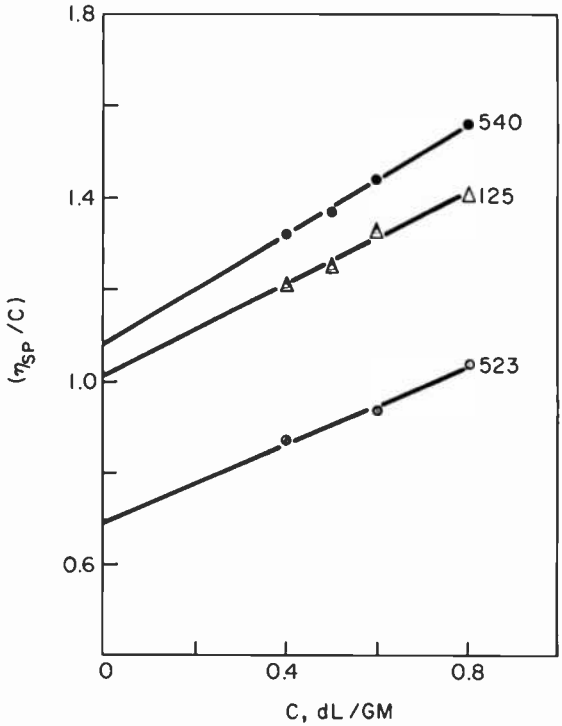


Fig. 6—Reduced viscosity vs concentration of PVA.

The effect of temperature on the intrinsic viscosity of PVA in water has been given by Sakurada[7] as

$$[\eta]_T = 1.07^{-(T-20)/10}[\eta]_{20°C} \quad [5]$$

It should be noted that the Elias equation is of the form $[\eta] = KM_w^a$, in which K and a are constants, usually independent of molecular weight but dependent on the nature of the polymer. This type of equation is usually referred to as the Mark-Houwink equation, and is applied to fractionated polymers or to polymers with narrow molecular weight distributions.

3.2 Basic Hydrodynamic Parameters

Flory[8] has evolved a hydrodynamic equation for the intrinsic viscosity of long-chain molecules in solution. In its simplified form,

Table 2—Hydrodynamic Parameters of Selected Samples of Poly(vinyl alcohol)[a]

PVA Type	Degree of Polym.	Degree of Hydration	$<r_o^2>^{1/2}$	$<s^2>^{1/2}$	R_h[b]	$[\eta]_{expt}$
540	2400	0.88	393Å	160Å	96Å	1.08
125	1800	0.996	336	137	82.	1.02
523	1800	0.88	307	125	75	0.69

[a] appropriate corrections in molecular weight made for residual acetyl groups.
[b] ζ is taken as 0.6 for the calculation of R_h. See Garvey, Tadros, and Wilson, *J. Colloid of Interface Sci.* 49 (1), 64 (1974).

this equation states that

$$[\eta] = \Phi \left(\frac{\overline{r_o^2}}{\overline{M}} \right)^{3/2} \overline{M}^{1/2}, \qquad [6]$$

where Φ is a fundamental constant, which for linear vinyl polymers is 2.1×10^{21}; $\overline{M}$ is the molecular weight, and $<r_o^2>$ is the mean square end-to-end distance of the unperturbed polymer chain. Thus, $(r_o^2)^{1/2}$ constitutes the root-mean-square end-to-end distance. Other basic hydrodynamic parameters* include $<s^2>^{1/2}$ radius of gyration, or the rms distance of the elements of the chain from its center of gravity, and R_h, the hydrodynamic radius.

The hydrodynamic parameters were calculated for PVA Samples 540, 125 and 523, and are given in Table 2.

It is interesting to compare our results for PVA 540, the material presently in use, to published values of other, similar, materials. Thus, Koopal and Lyklema[9] worked with Wacker (FRG)PVA of 0.88 degree of hydrolysis and molecular weight similar to our PVA 540. Their chain dimension parameters for this material are 380 Å for the end-to-end distance, and 160 Å for the radius of gyration, both in excellent agreement with our findings. As a check, the viscosity characteristic of PVA 540 at several concentrations were determined with a Ferranti direct-reading viscometer at various shear rates. The viscometer was first calibrated against NBS viscosity standards. Experimental values were converted to specific viscosities, which were then corrected to zero shear rate by extrapolation. Plots of corrected viscosities against concentration gave a value of 1.1 for the intrinsic viscosity of PVA 540, as against 1.08 determined by the generally accepted Ubbelohde method.

* See, for example, F. W. Bilmayer, Textbook of Polymer Science, Interscience, NY, 1964.

3.3 Rheology of Phosphor Slurry

We have investigated the rheology of our green phosphor model slurry, using a Ferranti-Couette viscometer, which has the advantage of being able to operate at varous shear rates, which is ideal for this examination. We found the viscosity of both resist and slurry to be Newtonian, i.e., following Eq. [1]. Results are shown in Fig. 7, with shear rate (sec^{-1}) plotted against shear force (dynes cm^{-2}). There was no evidence of thixotropy.

4. Pre-coat Properties

The role of the PVA pre-coat (PC) in screening has been a puzzling one for many years, although the necessity of a PC seems to have been well established. In our experiments with faceplate glass, we found the PC to be essential to good adherence. We therefore investigated some of the properties of the PC layer as they relate to adherence.

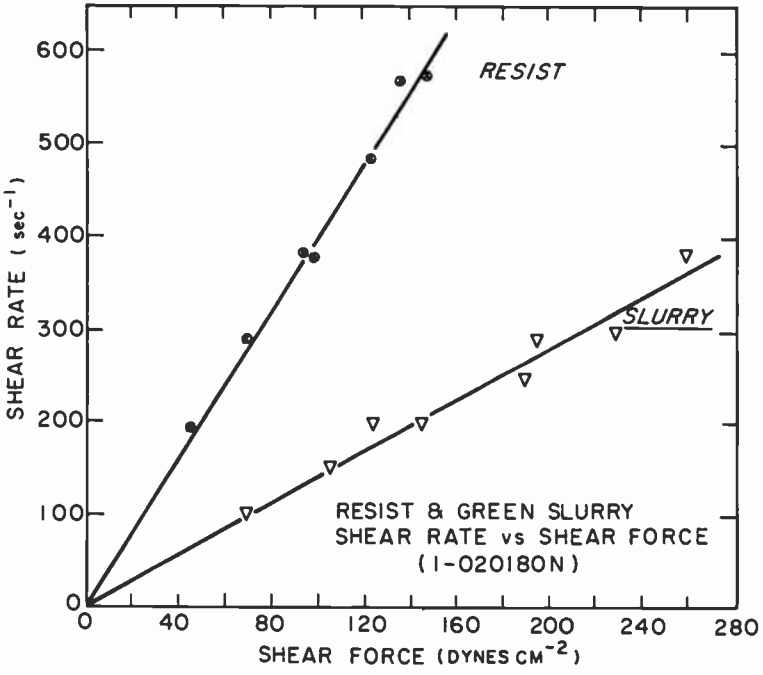


Fig. 7—Shear rate vs shear force for resist and green slurry.

4.1 Adherence and Pre-Coat Hydrodynamic Parameters

While it is difficult to correlate adherence with "normal" PVA pre-coat parameters, such as degree of polymerization and degree of hydrolysis, an excellent correlation is arrived at with hydrodynamic parameters. Fig. 8 is a plot of relative wet adherence against the hydrodynamic parameters derived in Table 1. We see that wet adherence increases with increasing hydrodynamic radius, radius of gyration, and the rms end-to-end distance of the unperturbed polymer chain. Thus, from Fig. 8, PVA 540 used as a pre-coat yields a phosphor screen with ten times greater wet adherence than PVA 523 and about 50% greater than PVA 125. It is of interest that in going from a degree of hydrolysis (DH) of 88% for PVA 523 to a DH of 99.6% for the PVA 125, both with the same degree of polymerization, the wet adherence increases seven-fold. It is apparent therefore that the hydrodynamic parameters of the PVA pre-coat are of

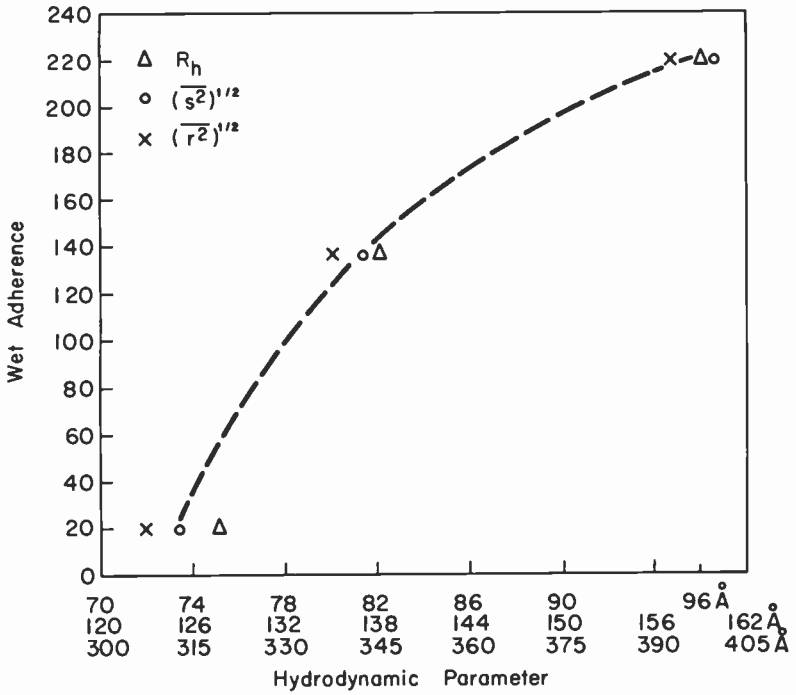


Fig. 8—Relative wet adherence vs hydrodynamic parameters.

great importance in determining the wet adherence of the phosphor screen.

4.2 Tacticity

Natta[12] devised the nomenclature which describes stereoregular configurations involving an asymmetric carbon. The *isotactic* configuration results when all the substituent groups R are above or below the plane of the main chain. If the R-groups lie alternately above and below the plane, the configuration is *syndiotactic,* and if the R-positions are random, the configuration is *atactic* or *heterotactic.* Fig. 9, from Finch[13], shows an example for PVA, where the polymer chain is shown in the fully-extended planar position.

Infrared spectra are used frequently to measure tacticity, and we have determined the tacticities for several of our PVA samples. Table 3, from Pritchard[10], after Liang and Learson[14], lists some characteristics of the IR spectrum of PVA. The bonds occurring at 916 cm^{-1} and at 850 cm^{-1} are characteristic skeletal vibrations of syndiotactic and isotactic PVA, and the percent tacticity is often arrived at from empirical relationships between tacticity and the ratios of the optical densities at these cited frequencies. For example, Kennedy and Willcockson[15] give:

$$\text{Percent syndiotactic} = 60\,(D_{916}/D_{850}) + 7 \qquad [7]$$

$$\text{Percent isotactic} = 59-78\,(D_{916}/D_{850}) \qquad [8]$$

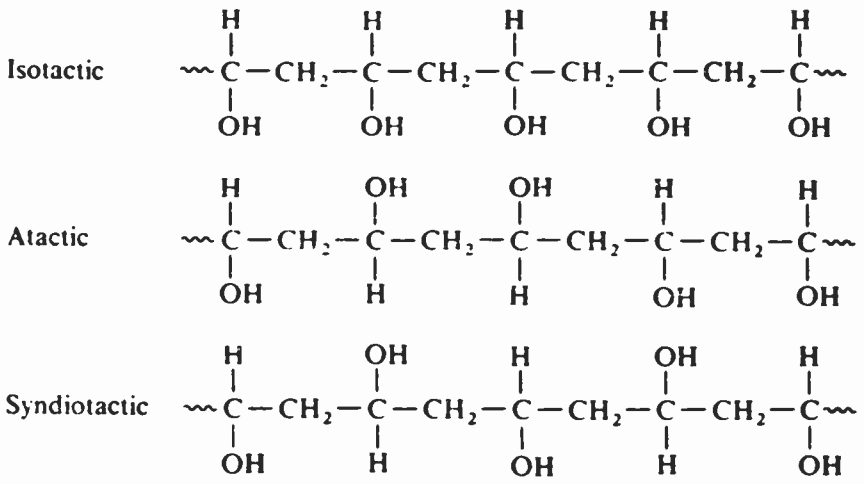


Fig. 9—Tacticity of PVA.

Table 3—Characteristics of the Infrared Spectrum of Poly(vinyl alcohol)[10,14]

Frequency (cm^{-1})	Relative intensity	Likely assignment
3340	Very strong	O—H stretching
2942	Strong	C—H stretching
2910	Strong	C—H stretching
2840	Shoulder	C—H stretching
1446	Strong	O—H and C—H bending
1430	Strong	CH2 bending
1376	Weak	CH2 wagging
1376	Medium	C—H and O—H bending
1320	Weak	C—H bending
1235	Weak	C—H wagging
1215	Very weak	
1144	Medium	C—C and C—O stretching
1096	Strong	C-O stretch & O-H bend
1087	Shoulder	
1040	Shoulder	
916	Medium	Skeletal
890	Very weak	
850	Medium	Skeletal
825	Shoulder	CH2 rocking
640	Medium, very broad	O—H twisting
610	Weak	
480	Weak	
410	Weak	
360	Shoulder	
185	Very weak	
135	Very weak	

Table 4 lists the tacticities and carbonyl for several of our PVA samples. It should be noted that (1) the sum of the tacticities is not exactly 100% due to the formulae utilized; and (2) the tacticities of the presently used pre-coat, PVA 540, are not greatly different from those of PVA 125 and PVA 523. Differences are found in syndiotactic and isotactic content compared to PVA 425. It is interesting that PVA 125, nearly 100% hydrolyzed, showed no detectable carbonyl; PVA 425 (about 96% DH) had 0.8% carbonyl; while PVA 523 and PVA 540 each 88% DH), had about 4% carbonyl. Carbonyl content varies inversely with DH.

Table 4—Tacticities and Carbonyl for Several PVA Samples

PVA	D.H.	Syndiotacticity	ISO.	Hetero.	Carbonyl
125	99.6%	25%	36%	40%	0
425	95.5%	20%	42%	38%	0.8
523	88%	25%	36%	40%	4.0
540	88%	25%	37%	38%	3.7

4.3 Pre-Coat Hydroxyl Content

Results obtained on PVA powders, using FT-IR, 100 scans, showed a relationship between wet adherence and the hydroxyl concentration of the pre-coat PVA. This is shown in Fig. 10 for green phosphor model-slurry where three different pre-coats were used, all at 0.5% concentration and all at pH 2.7. These results corroborate the findings that for this PVA group of low (79,000) molecular weight, wet adherence was found to increase with increasing degree of hydrolysis. It should be noted, however, that the molecular weight of the PVA appears to have an over-riding effect on wet adherence. Thus, the 88%-hydrolyzed PVA 540 showed a wet adherence nearly an order greater than the corresponding low molecular weight PVA. Again, the important aspects are the hydrodynamic parameters.

It is interesting to compare the hydroxyl concentration results with those for wet adherence as a function of degree of hydrolysis, as shown in Fig. 11, for 0.3% PVA. We see, as expected, the sharp dependence of wet adherence with degree of hydrolysis for these three samples of PVA. Note that the lowest adherence, 22 in^{-2}, is

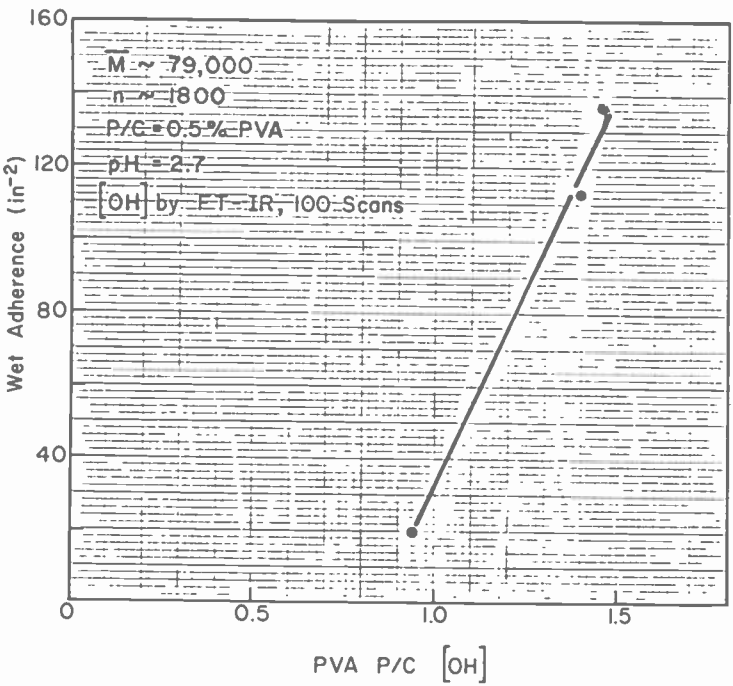


Fig. 10—Wet adherence vs pre-coat hydroxyl concentration.

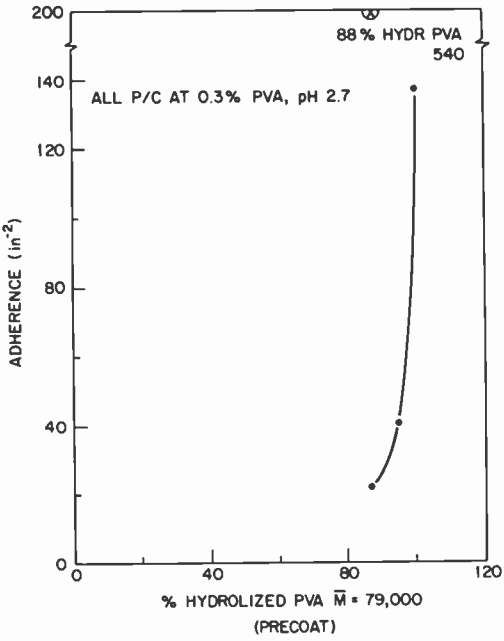


Fig. 11—Wet adherence vs degree of hydrolysis of PC-PVA.

for PVA 523 with DH = 0.88 and DP = 1800. For PVA 540, with DH = 0.88, but with DP = 2400, the adherence was about 200 in^{-2}, as indicated at the top of the figure.

4.4 Adherence and Pre-Coat pH

Wet adherence has been found by Harper[16] and by Wilcox[17] to be affected by pH. This effect was corroborated here, and was also found to depend on the type of glass substrate. Samples of model slurry were prepared, and the wet adherence was checked for 0.5% PVA at various pH's, using FPG or soda lime glass (SLG) substrates.

It is seen from Fig. 12 that wet adherence for FPG substrates increases with decreasing pre-coat pH, the effect leveling-off as pH = 1 is reached. The screen *quality* is poorer at the very low pH pre-coats, however, and we chose pH = 2.7 for our pre-coat work. The adherence of 4 screens made on SLG shows better adherence than those on FPG; adherence levels off at a pH of about 3.7 and remains constant to at least pH = 1.5. Other differences due to glass type will be described in the section on glasses.

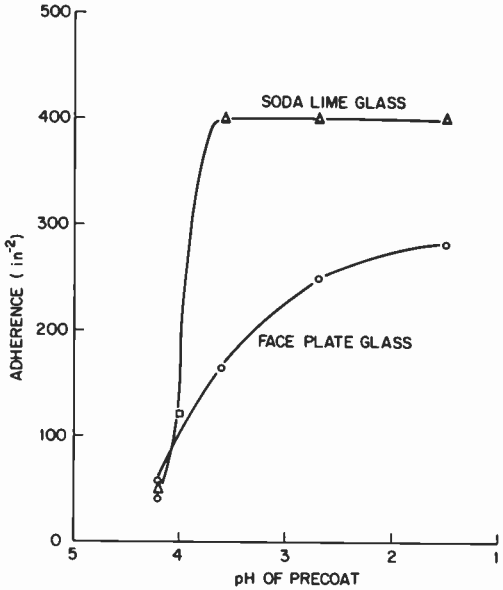


Fig. 12—Wet adherence vs pre-coat pH.

4.5 Pre-Coat Thickness Effects

A series of screens was prepared where the pre-coat thickness was varied by using PVA solutions of different concentrations, all at pH 2.7. Tallysurf measurements[18] were made of the PVA pre-coat thickness, and a plot of PVA concentrations versus PC thickness is shown in Fig. 13 for 0–4% PVA. The relative adherence plotted against PVA pre-coat thickness, for thicknesses from 0.18 μm to about 1.5 μm is shown in Fig. 14A. There is a striking decrease in adherence in going from 1800 Å to 2000 Å, followed by a constant adherence to at least 1.5 μm thickness of pre-coat. Further investigations of even lower concentrations of pre-coat PVA yielded results shown in Fig. 14B. Here, we have plots of wet adherence versus percent PVA concentration in the PC for SLG and FPG. The optimum PVA concentration was about 0.3%, which, for FPG, yielded a PC thickness of about 600 Å. Thickness points of 400 Å and 1000 Å are also indicated. Again, the SLG substrates show greater adherence. Knowing the area, density of solid PVA, and the volume (0.08 nm^3) and the cross-sectional area of a PVA segment (0.2 nm^2) from the work of Koopal and Lyklema[9], it can be calculated that the optimum pre-coat thickness consisted of about 200

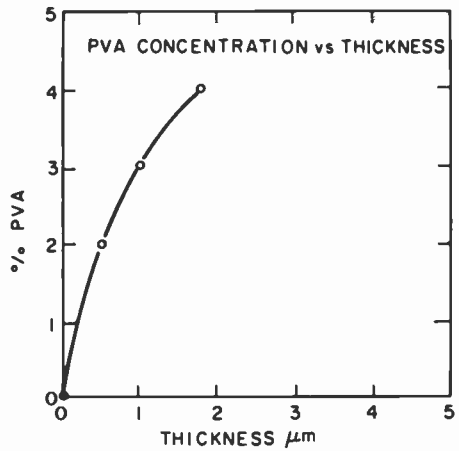


Fig. 13—PVA concentration vs PC thickness.

monolayers. Our work also indicated a retention of 3.9% for PVA pre-coat after spin-off. Apparently, there is an optimum pre-coat thickness for maximum adherence, again involving interfaces 1 and 2. Yakhinin[20] has reported that forming an adsorbed layer of PVA

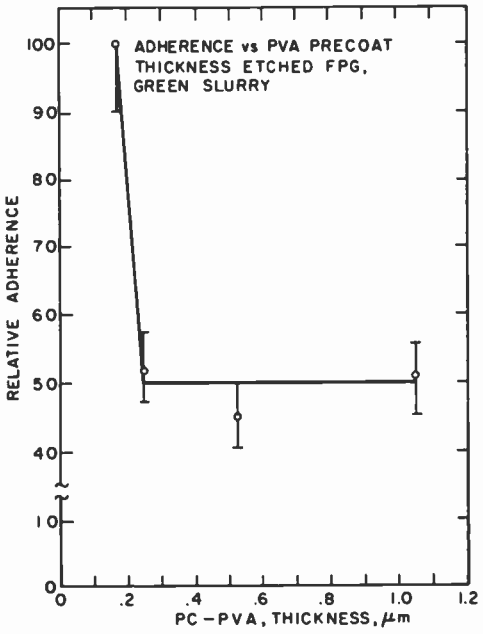


Fig. 14A—Relative adherence vs PC thickness.

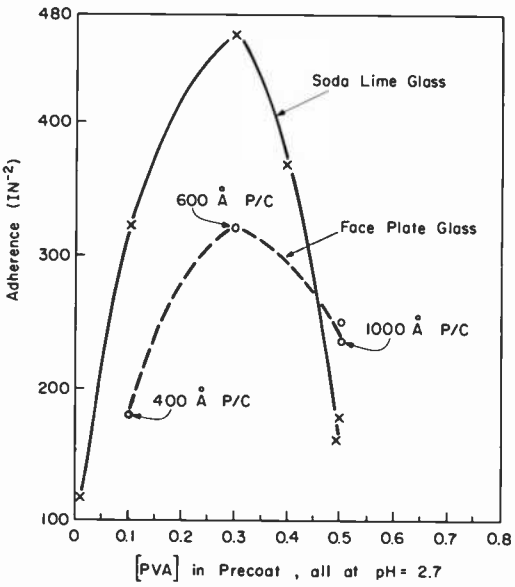


Fig. 14B—Wet adherence vs pre-coat PVA concentration.

decreased the contact interaction between a glass plate and 45 μm particles of rutile.

4.6 ζ-Potential

A potential difference exists between the surface of a particle and the bulk solution because of a charge density localized in the plane of its surface. Since the system as a whole is electrically neutral, this surface charge is balanced by an opposite equal excess ionic charge in the liquid phase that will concentrate in the vicinity of the particle surface.

When such a system is put in an electric field, the particles move toward the oppositely charged electrode, the counter ions moving in the reverse direction. The particle velocity increases with the ζ-potential of the particle. Thus, by using an electrophoretic mobility apparatus, we can derive ζ-potentials.

The greater the ζ-potential of a particle, the greater the repulsive force between it and similar particles. If two such high ζ particles are driven together thermally or mechanically, they will tend to separate in spite of van der Waal's forces tending to bind them

together. This could conceivably be of importance in such screening characteristics as porosity, line definition and cross-contamination.

The Micromeritics Model 1202 Electrophoretic Mass Transport Analyzer is capable of measurements under very high solids loading conditions. With this instrument, we determine the average particle electrophoretic mobility:

$$V_\epsilon = \frac{\Delta W \lambda}{ti\,\phi(1 - \phi)(\rho_p - \rho_L)}, \qquad [9]$$

where V_ϵ is the mobility, ΔW is the mass difference, i is the current, t the time, λ the specific conductance, ϕ the volume fraction of dispersed matter, and ρ is the density of the particle (p), or of the liquid (L).

The zeta-potential is then calculated from the Smoluchowski equation relating ζ to electrophoretic mobility, viscosity η, and dielectric constant D_t:

$$\zeta = \sigma\left(\frac{4\pi\eta}{D_t}\right) \qquad [10]$$

or, in the case of our instrument,

$$\zeta = \frac{4\pi(\Delta W)\lambda\eta}{ti\,\phi\,(\rho_p - \rho_L)D_T}. \qquad [11]$$

Since our 3 × 3 inch FPG slides could not be accommodated in the apparatus, faceplate glass was crushed and ground to a powder. Blank FPG had $\zeta = +4$ to $+6$ mV. However, FPG which had been etched (5% NH_4F) and washed had $\zeta = +130$ mV, while etched, washed, and PVA-coated FPG had $\zeta = +73$ mV. This was the first indication that the adsorbed ions due to etching of the glass imparted a highly positive zeta-potential to the FPG, which is reduced but remains highly positive after adsorption of small amounts of PVA from the 0.5% pre-coat. Fig. 15 illustrates the effect on the zeta-potential of increasing amounts of PVA. While this particular curve was obtained for pre-pigmented red-emitting phosphor, similar results were obtained with green. This Y919F Fe_2O_3-coated red phosphor has a specific surface area of 0.350 m^2gm^{-1}; the total phosphor area was 15.75 m^2.

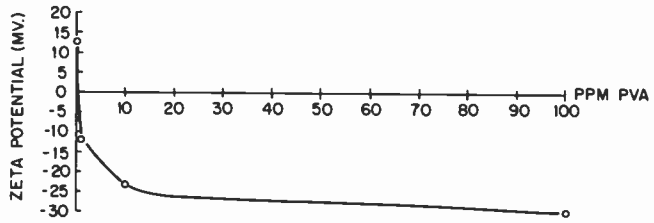


Fig. 15—Zeta potential vs PVA concentration.

Tadros[21] has published on the variation of zeta-potential with concentration of PVA adsorbed on an organic substrate and has obtained similar shapes, which he explains by invoking the macromolecules having a flat configuration which leads to thicker adsorbed layers. This will cause a shift in the shear plane toward the solution, leading to a lowering of zeta-potential until it reaches a limiting value.

Koopal and Lyklema[9] investigated the adsorption of PVA on silver iodide by double layer measurements using PVA samples similar to PVA 540. They report that the first 0.15 mg/m^2 adsorb in trains or flat loops; PVA in excess of this amount is adsorbed as loops. A detailed analysis of this first layer found there were twice as many trains as loops.

Garvey et al.[23] have compared different adsorbed layer thicknesses of PVA on polystyrene latex particles. For a 0.87 hydrolyzed PVA (M_w = 67,000, $\langle\overline{S}^2\rangle$ = 117 Å, R_h = 67.5 Å), they report an effective thickness of the adsorbed PVA layer of 380 + 50 Å. It is interesting to compare this value for a PVA with small hydrodynamic factors and molecular weight with our value of 600 Å for the optimum PVA pre-coat thickness for adherence.

4.7 Adsorbed PVA

Examination of the green phosphor slurried in PVA showed, by the boric acid-iodine method of Zwick[19], that PVA *is* adsorbed onto the phosphor particles. This analytical method, highly sensitive for small amounts of PVA, was extended to larger amounts. A typical calibration curve is shown as Fig. 16, and Fig. 17 shows an adsorption curve for PVA in pre-pigmented red-emitting phosphor; axes are given in γ-PVA per millimole Y_2O_2S:Eu, and γ-PVA per square

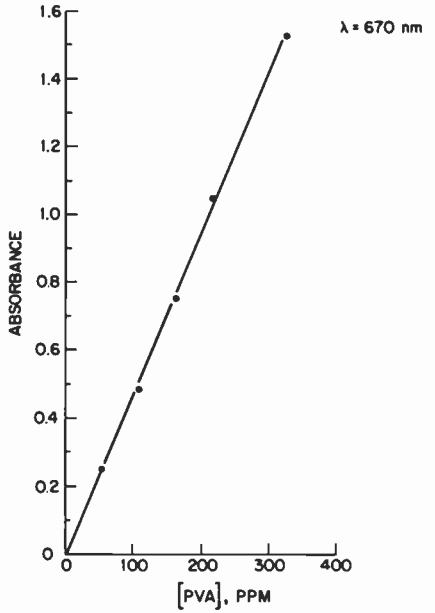


Fig. 16—Calibration curve for PVA adsorption.

meter surface area of phosphor. Similar results were obtained for green phosphor.

While one may speculate about a negative-ζ particle coated with PVA impinging on a positive-ζ FPG substrate to provide initial sticking, our measurements indicate that the FPG is even more positive *without* the pre-coat. However, our adherence tests show adherence failure without the pre-coat. Zeta-potentials by themselves, therefore, do not provide a complete explanation, but do constitute an important aspect of adherence, particularly with respect to the FPG.

In polymers and in glass, adhesion is believed to involve functional groups with large group dipole moments.[2] One could then have good adherence due to the asymmetrical electron density distribution in the contact zone, leading to an electrical-double-layer. Table 5 gives some pertinent values for several materials.

In the reactions of organic molecules with the molecules of a substrate, the major factors are the electronic processes that occur in the molecules and in the contact zone. If an asymmetric electron

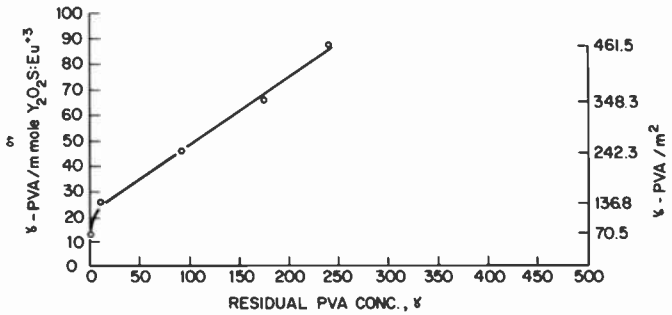


Fig. 17—Adsorption of PVA on pre-pigmented red phosphor.

density distribution occurs, with a double layer being formed, then the charge density σ can be defined as

$$\sigma = en, \qquad [12]$$

where n is the number of donor-acceptor pairs. The electron negativity series for functional groups in order of decreasing electron-donor properties is

Donor
NH_2>OH>OR>OCOR>CH_3>C_6H_3>halogens>COOR>CO>CN
Acceptor.

Usually, the most reactive group in a polymer is the *hydroxyl.*

5. Glass Substrates

It became apparent during the course of our investigations that the nature of the glass substrate was of great importance to the adher-

Table 5—Group and Total Dipole Moments[2]

Polymer	M_w	Group Dipole Moment ($\times 10^{-18}$)	Total Dipole Moment ($\times 10^{-18}$)
Vinyl acetate	86	1.75	1.75
Ethyl acetate	88	1.86	1.86
Polyvinyl acetate	24,000	1.71	28.7
	60,000	1.68	44.4
Polystyrene	12,900	0.08	0.89
	42,900	0.09	1.83

ence process. Therefore, some work was done on glasses, particularly face-plate glass.

Glass is generally thought of as an amorphous material, having a random network structure[23], pure silica being a continuous, random, three-dimensional polymer. Glasses also contain additives (impurities) that seem to reduce the melting point and viscosity, to control expansion, improve durability, and to prevent crystallization. Our FPG also contains lead for x-ray protection. The nature of the surface of a glass is controlled by the melt temperatures, volatiles lost during the melting step, volatiles lost during the hot-forming step, water vapor absorbed during forming and cooling, hydrocarbons from lubricating oils, and impurities from dies during the pressing cycle.

As discussed in previous sections, we noticed a difference in adherence between soda lime glass and faceplate glass. Since SLG could not be used in our CTV picture tube process, we concentrated on FPG.

5.1 Surface Hydroxyl (Silanol) Groups

It is well known that glass has surface hydroxyl groups.[23] The concentration of such groups can be increased by strong acids or hot strong ammonia. That such silanol groups are involved in adherence was shown by two experiments:

(1) FPG was washed in hot, concentrated ammonia, followed by the normal wash, pre-coat, slurry and developing. The adherence was found to be equivalent to the HF-etched control.

(2) FPG was etched in the normal fashion, after which surface silanol groups were removed by hexamethyldisilazine (HMDS), a well known method for removing surface hydroxyl groups.[25] In this method,

$$(CH_3)\ SiNHSi(CH_3)_3 + ROH \rightarrow ROSi(CH_3)_3, \qquad [13]$$

so that the surface hydroxyl is converted efficiently to trimethylsilyl groups.[26] Complete screens made in this fashion showed, on developing, excellent lateral coherence but no adherence to the FPG, sliding off the glass as a complete coherent screen. We find then that the wet adherence of screens is related to the glass silanols.

There is general agreement in the literature on glass that there are spectroscopic changes as glass is dehydrated and re-hydrated.

Thus, if a glass is heated above 450°C, the hydration process is not easily reversible, and the reaction is:

$$\begin{array}{l} \text{OH OH} \qquad\qquad\quad \text{O} \\ \quad | \quad | \quad T > 400°C \quad \wedge \\ \text{Si-O-Si} \longrightarrow \text{Si-O-Si} \end{array} \qquad [14]$$

where the OH's are so-called vicinal hydroxyl groups.

Fig. 18 indicates the possible mechanisms involved in etching and pre-coating. Part (a) shows dehydrated vicinal surface hydroxyls. In part (b), the formation of silanol groups on the surface is due to etching. Part (c) shows the pre-coat PVA layer applied to the glass surface, and Part (d) indicates the oxygen-bridge bonding of pre-coat to glass surface.

This figure does not take into account the ions leached from the glass by the etchant and the interfacial results cited later in this report.

5.2 Leaching Impurities from Faceplate Glass

Faceplate glass slides, cut from the same faceplate, were leached

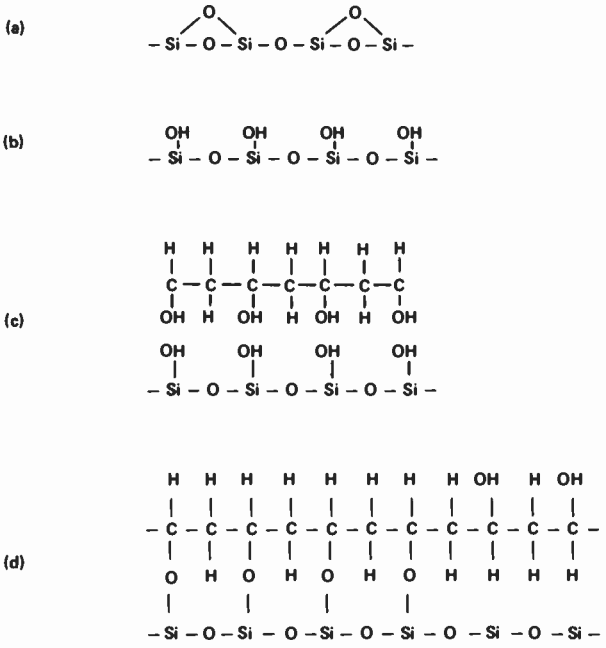


Fig. 18—Model for etching and pre-coating.

overnight in triple-distilled water in plastic containers for 0.5, 1.0, 2.0, and 4.0 hours. Runs included blanks, unwashed FPG, washed and etchd FPG, and washed, etched, and pre-coated (0.5% PVA) FPG. Samples of the water were then analyzed by optical emission spectrometry and atomic absorption.[24] The results are summarized in Table 6.

We see from Table 6 that etching serves to release a number of ions from FPG and that the concentration leached out after etching is time-dependent. The effect of the pre-coat is interesting in that pre-coating after etching results in lower leached ion concentrations in every case, even at the high (4 hour), leach time. Silicon is striking in that it is absent in the etched-precoated FPG but present in the etched FPG. It is felt that the leached ions are partially complexed by the PVA and silanols thus resulting in lower concentrations in the leach water. Relz et al.[27] have reported on the fixation of metal complexes on glasses.

Results of our experiments indicate the possibility that glass impurity ions are also involved in adherence. A possible model to explain our results at this point is shown in Fig. 19. Here, FPG has been etched to give surface silanol groups and impurity ions M. These impurity ions could be coordinated betwen glass silanols and PVA hydroxyls of the pre-coat at interface I_1 (see Fig. 2). Other impurity ions could then coordinate between hydroxyls of pre-coat PVA and of phosphor particle PVA, as in interface I_2 of Fig. 2, thus resulting in screen adherence. As will be discussed later, this is not the complete model.

5.3 Interfaces

For the exploration of interfaces (particularly I_1 and I_2) as they relate to tube processing steps in particular, we chose to investigate

Table 6—Leachable Impurities From Faceplate Glass and the Effect of Etching and Pre-Coating

FPG		Unwashed				Washed/Etched				Wash/Etch/Precoat			
Time(hrs)	Blank	0.5	1	2	4	0.5	1	2	4	0.5	1	2	4
Ba(ppm)						1	—	10	10	0.3	0.3	1	2
Sr(ppm)						12	12	60	60	10	10	12	12
Si(ppm)					2	2	2	10	6	—	—	—	—
Mg(ppm)	0.3	0.3	0.3	0.3	2	100	100	300	300	2	10	50	100
Ca(ppm)						60	60	300	300	15	15	100	100
Al(ppm)						6	6	120	120	6	6	30	30
Pb(ppm)						20	20	200	200	—	—	50	50

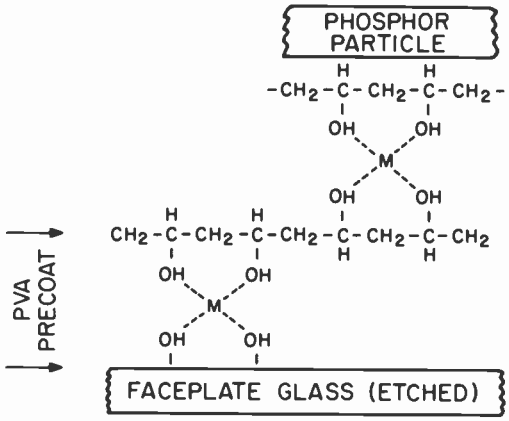


Fig. 19—Impurity ions in adherence.

contact angles. A commercial type Roué-Hart contact angle instrument was used, where advancing and receding angles could be studied as a function of tilt angle. Specially-prepared water was used as the fluid, and drops were obtained reproducibly from a calibrated Teflon microsyringe.

Contact Angles

More than 150 years ago, Young proposed treating the contact angle of a liquid on a solid as a result of the equilibrium of the drop under three surface tensions, the liquid-vapor interface (γ_{LV}), the solid-liquid interface (γ_{SV}), and the solid-vapor interface (γ_{SV}), as shown in Fig. 20. The Young equation is

$$\gamma_{SV} - \gamma_{SL} = \gamma_{LV} \cos \theta, \qquad [15]$$

where θ is the contact angle.

Since glass is a "high energy surface" (i.e., high specific surface free energy) and organics have low specific surface free energies, organics would be expected to spread on glass, as there would be a large decrease in the surface free energy of the system. We can thus consider the interfaces of Fig. 2 as a high-energy surface (FPG), on which is a low-energy surface (pre-coat), on which is another low-energy surface (phosphor PVA), which, in turn, has been deposited on a high-energy surface (phosphor).

The critical surface tension, γ_c, of PVA solid is 37 dynes cm^{-1} as reported by Scholz et al.[28], who also found γ_c to be between 40–45

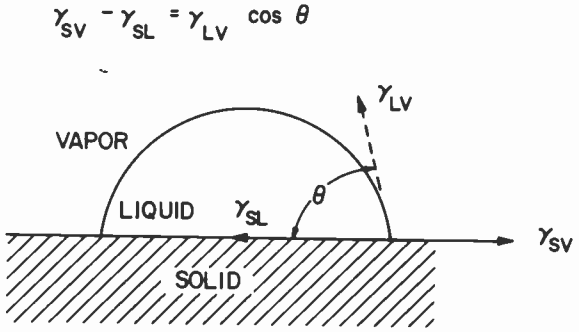


Fig. 20—Contact angles of a sessile drop.

dynes cm^{-1} for a series of hydroxyl-rich surfaces of the starch-polymer type. Spreading on low energy surfaces is caused by the lowering of the surface tension of water. Since γ_c of solid PVA is in the neighborhood of 37 dynes cm^{-1}, the solution will spread on the surface of PVA whenever a wetting agent lowers the surface tension of water below that level.

Processing and Contact Angle

Faceplate glass was taken through the steps of washing, etching, and pre-coating with 0.5% PVA 540, and contact angles were determined for each step at various tilt angles. The results were independent of tilt angle, and are shown in Fig. 21. It is seen that washed FPG has a contact angle of 30°. Etching the FPG decreased the contact angle to 10°, and pre-coating increased the contact angle to 28°. Continuing the same FPG through another wash, etch, pre-coat cycle, we see that the second wash increased the contact angle to 70°, the subsequent etch decreased the contact angle to 29°, and the following pre-coat and wash increased the contact angle again to 75°. Although only two processing cycles are shown in the figure, additional processing resulted in a repetition of these cycles. It is apparent that etching decreases the contact angle greatly, which is to be expected if the FPG surface is hydroxylated, and thus, becomes more hydrophilic. Pre-coating with PVA increases the contact angle, making the glass less hydrophilic.

The first three steps of Fig. 21 correspond to the initial processing for the first phosphor in a CTV kinescope. This first phosphor slurry is applied, dried, exposed, and developed. At this point, however,

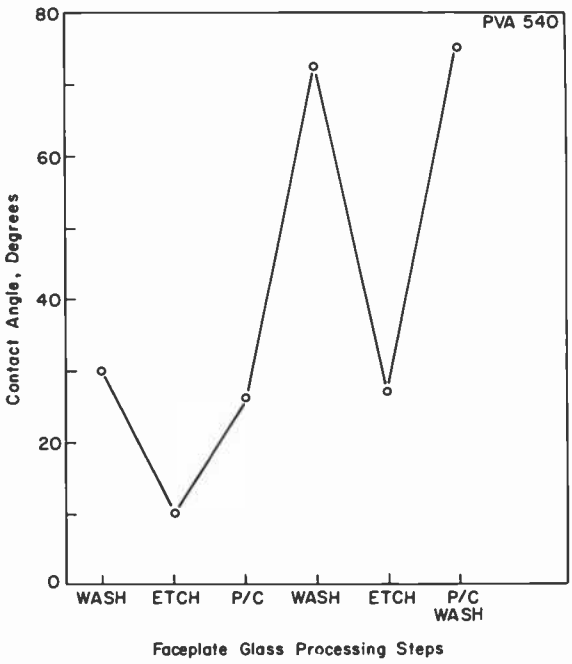


Fig. 21—Contact angle vs processing step.

the FPG contact angle increases toward hydrophobicity (before the second phosphor slurry is applied). What happens to adherence? We investigated this by making blue screens and measuring the adherences for FPG treated to have $\theta = 30°$ and $\theta = 70°$. We found adherence = 220 for $\theta = 30°$, and 156 for $\theta = 70°$, or a ratio of 1.41. The work of adherence on a surface, W_a, is given by

$$W_a = \gamma_L (1 + \cos\theta) \qquad [16]$$

γ_L was measured with a DuNouy tensiometer; $W_a^{30°}$ was calculated as 125.1 and $W_a^{70°}$ as 90.1, a ratio of 1.39. It would appear therefore that adherence of our phosphor screens is related to the work of adherence on the surface, which would therefore call for the smallest contact angle, which we obtain by etching the FPG.

When we wash the pre-coat layer, θ increases to 70° instead of decreasing to 10°, as it did after the first etching. This is interpretable as further evidence that a hydrophobic layer exists in interface 1, quite possibly contiguous to the glass as shown in Fig. 22. It may be formed by a coordinated metal-ion-PVA reaction, and there is

unreacted PVA on its surface. Washing the surface will remove interface I_{1B} (Fig. 22), thereby exposing interface I_{1A}. Subsequent etching might remove most, but not all, of 1A, thereby decreasing θ, but not to the θ value of the original etched surface. It is of interest that recent work has appeared on metal in complexes with PVA.[29] In terms of adhesion, partly-hydrolized PVA has been described for highly hydrophobic polymer surfaces.[30]

6. Adherence as a Function of Processing Steps

Our technique of measuring adherence by jet-impingement 3 × 3 inch sections of phosphor screens on FPG enables us to examine the variation in adherence during the processing procedure.

The adherence of dried pre-coat was measured by precoating a slide, drying it, and jet-impinging. The hole made by the jet is rendered visible by treating the slide with methylene-blue-dye dissolved in methanol. This dye stains the PVA so that the hole is easily measured and the adherence calculated. Similar adherence measurements were made after the slurry was applied and dried, after exposure, and after development. The results are shown in Fig. 23. We see that pre-coat adherence is very low, and dried slurry adherence is only about 35% higher. Exposing the screen increases the adherence by a factor of fifty, but the developing step reduces it by a factor of three. Soda lime glass, while showing identical adherence for pre-coat and slurry, shows better adherence than FPG after exposure and after development.

These results indicate the importance of interfaces I_3 and I_4 in the adherence procedure. For these cases, the dichromate sensitization step becomes important in cross-linking the PVA and thus bonding the phosphor particles. Grimm et al.[30] have discussed the photochemical reactions in a dichromated resist, and this aspect will

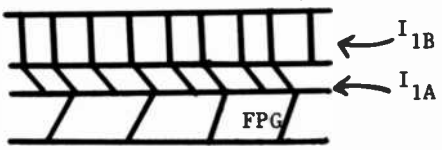


Fig. 22—Coordinated metal-ion interface.

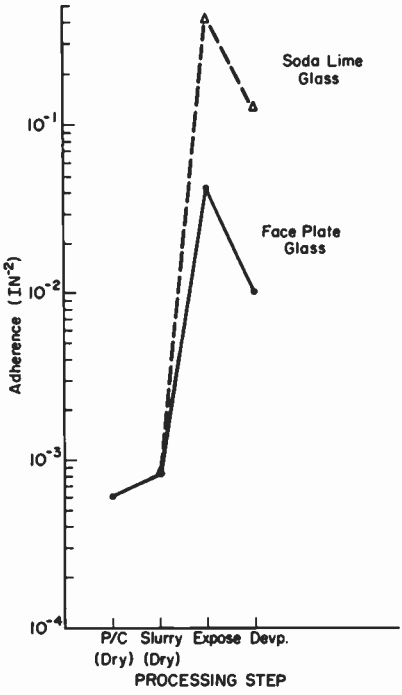


Fig. 23—Adherence vs processing steps.

not be reviewed here. From our point of view, the sensitized PVA-phosphor slurry provides for inter-particle bonding, the first particles of which form I_2 (Fig. 2).

7. Contact Characteristics of Model Slurries

The contact properties of model slurries, prepared with PVA540, were investigated since these properties are important in commercial tube production. Tube simulation studies were carried out using FPG, which was washed, etched, pre-coated with PVA540, and dried.* Contact angles were determined on this surface for water and for pre-pigmented blue phosphor model slurry (surface tension = 40.2 dynes cm^{-1}), which corresponds to the deposition of the first phosphor. This FPG substrate was then water-washed, corre-

* The abbreviated notation developed during this work would be: FPG/W/E/PC540/D//.

Table 7—Blue Model Slurry Contact Properties

Processing Steps	Blue Slurry		H_2O	
	θ	W_A	θ	W_A
FPG/W/E/PC/D//	58° ± 3°	62	37° ± 3°	121
FPG/W/E/PC/D/W/D//	39° ± 3°	72	66° ± 3°	94

sponding to the developing step, and dried, giving a total history of FPG/W/E/PC540/D/W/D//, and contact angles were again determined for water and for the blue slurry, which now corresponded to the *second* phosphor layer.

Consistent with the results shown in Fig. 21, the surface seen by the second layer of blue phosphor is more hydrophobic than that seen by the first layer of phosphor, so that the energy of wetting of the blue phosphor is different, depending on whether it is deposited as a first layer or as a second layer. Similarly, W_A, the work needed to remove a film, increases as the surface becomes more hydrophobic. These results are summarized in Table 7. Interestingly, the third phosphor deposition "sees" a more hydrophilic surface, perhaps due to build-up of PVA on the surface.

It would appear from these results that the more hydrophilic a surface, the more "organophobic"; and the more hydrophobic the surface, the more "organophilic" as applying to the model phosphor slurry.

Having investigated the contact angle of (blue) slurry, we examined separately deposition of the resist component (PVA540) on pre-coats of PVA540 and PVA165. A comparison of the results shown in Table 8 with Table 7 shows (a) PVA540 resist acts in an identical fashion for contact angle, with PVA540 or PVA165 pre-coats and (b) that the contact angle is determined primarily by the resist portion of the slurry. The surface tension of the resist was measured as 39 dynes cm^{-1}.

Results shown in Table 7 indicate that the PVA pre-coated surface is hydrophilic, and organophobic. When the PVA pre-coated

Table 8—Contact Angles of PVA540 Resist on PVA540 and PVA165 Pre-coats

Processing Step	PVA540PC θ	PVA165PC θ
FPG/W/D//	46° ± 3°	42° ± 3°
FPG/W/E/D//	40° ± 3°	45° ± 3°
FPG/W/E/PC/D//	59° ± 3°	61° ± 3°
FPG/W/E/PC/D/W/D//	45° ± 3°	42° ± 3°

surface is washed, the new surface becomes more hydrophobic and more organophilic to the blue slurry. While this might be due to acetyl groups on the PVA540 pre-coat, the results in Table 8 for PVA540 and for PVA165 show that this is not the case.

What would be the effect of changing the type of PVA pre-coat on the contact angle of water? For these experiments, PVA165, D.H. = 99.7%, was used as the pre-coat. Results, including the previous PVA540 pre-coat for comparison purposes, are given in Table 9.

It is interesting to note that for PVA165 the ratio of W_A is 0.9, and the adherence ratio is 1. Similarly, for PVA540, the W_A ratio is 1.3 and the adherence ratio is 1.4.

8. Other Effects

FPG Surface Additives

Since the etching step appears to have two functions, silanol formation and release of impurity ions, experiments were performed where the FPG surface was exposed to a given ion and dried prior to pre-coating, thus eliminating the etch step. Fair adherences were obtained in this manner with an apparent correlation between adherence and the solubility product of the metal-ion-sulfide. That the green (sulfide) phosphor was necessary for adherence was established by applying resist (no phosphor) to the surface-treated unetched FPG. No adherence was obtained. With green slurry, adherence appeared. For the ions tried, the adherence order was: copper (S.P. of the sulfide 10^{-47}) best, followed by cadmium (S.P. 10^{-29}) and zinc (S.P. 10^{-23}).

Pre-Coat Additives

We also investigated the addition of inorganic ions to the PVA pre-coat to provide a highly cross-linkable region at interface I_1, the critical interface for adherence. Y^{3+} was used as the additive, since previous work here[31] had shown it to have a cross-linking effect.

Table 9—PVA165 and PVA540 Pre-coats

	PVA165			PVA540		
	θ_{H_2O}	W_A	ADH.	H_2^0O	W_A	ADH.
FPG/W/E/PC/DRY//	52°	108	115	37°	121	222
FPG/W/E/PC/DRY/W/D//	40	119	115	66	94	156

Experimental results did indeed show an increased adherence for screens prepared with yttrium-doped pre-coats. Fig. 24 shows the yttrium-effect on the adherence of our green model slurry, the optimum being at about 50 micrograms of yttrium. However, this effect was absent with slurry containing triethyleneglycol because the yttrium evidently complexed with the triethyleneglycol in the slurry, thereby becoming unavailable for crosslinking at the interface.

Slurry Additives

Triethyleneglycol (TEG) was investigated as an additive to our green model slurry. Using our standard procedure, namely, FPG/W/E/PC/D//, we found that the adherence increased monotonically with increasing TEG, as shown in Fig. 25. It is believed that this particular effect can be accounted for by the increase in sensitivity found with the addition of TEG. Over the range of TEG/PVA covered in our experiments, Fonger[32] has reported a monotonic decrease in threshold, τ, also shown in Fig. 25. This increase in sensitivity produces an increase in adherence for a constant time (and intensity) of exposure.

Aging Effects in Slurry

During the course of these investigations, it was observed that

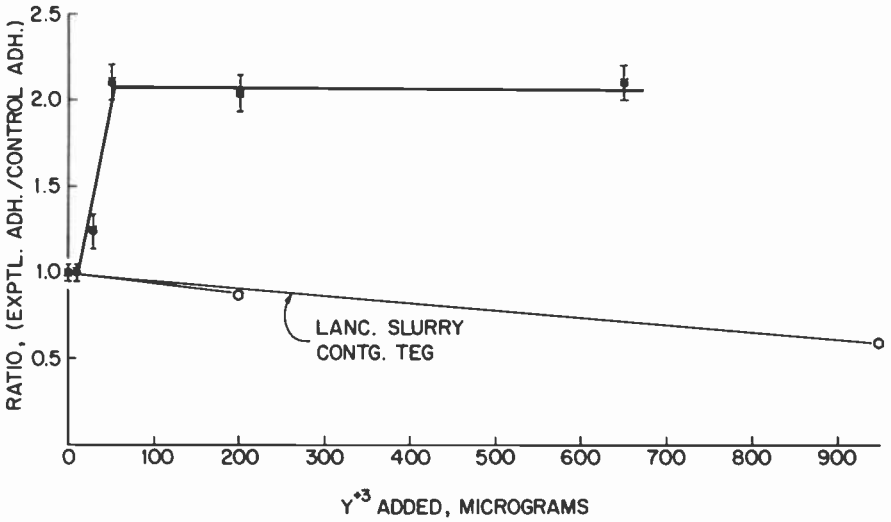


Fig. 24—Adherence vs pre-coat Y^{+3} concentration.

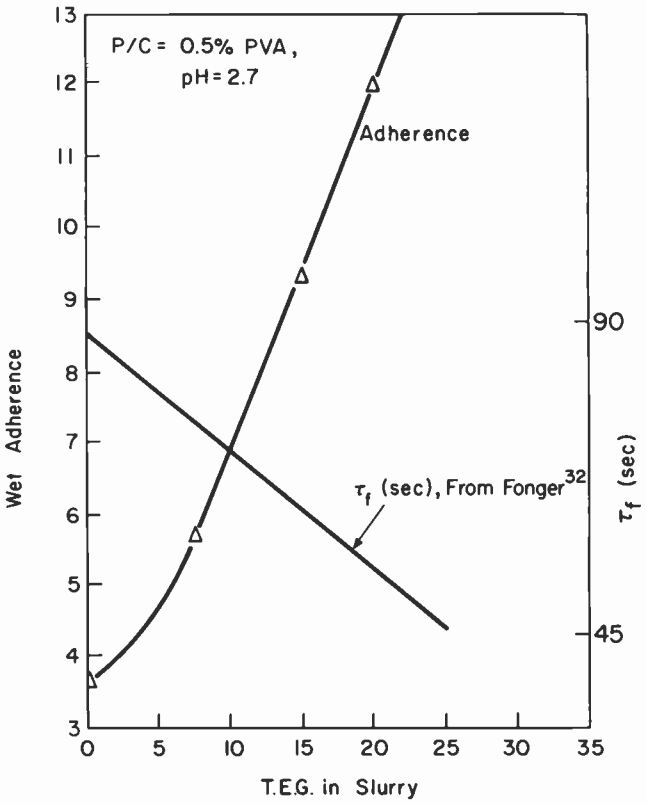


Fig. 25—Effect of TEG on adherence.

aging effects were taking place, i.e., adherence appeared to improve as our green slurry aged. This was examined in some detail, and the results are shown in Fig. 26 where we see that maximum wet adherence occurred after about six days aging. Such effects might be related to cluster growth in the slurry to an optimum size. Caution must be observed in investigating aging effects, since other parameters may be operative. For example, aging in blue slurry showed a decrease in adherence that could be correlated with the effect of Mg^{++} ions (used in the phosphor) on the hydroxyl ion concentration. This is shown in Fig. 27.

Plasma-Treated Pre-Coats

Plasma-treating the pre-coat could effect cross-linking. It might also effect the aldehyde and ketone concentrations in the PVA and the

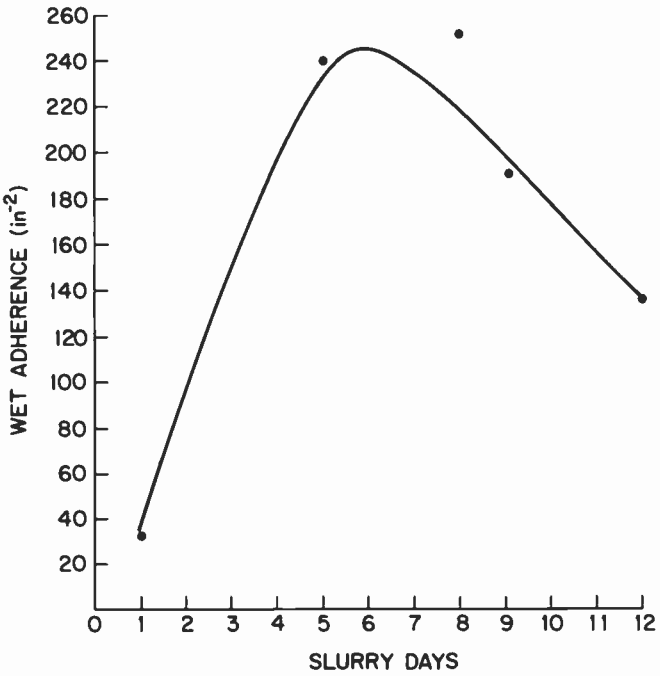


Fig. 26—Adherence vs aging for green slurry.

proportion of high-molecular-weight polymer. Exploratory runs were made[33] on PVA-precoated faceplate glass. Results are summarized in Table 10.

In summary, a one-minute treatment with Ar or Ar + H_2 plasma had no effect. However, extending the plasma time to ten minutes resulted in a large decrease in adherence for Ar-plasma, but in an increase for samples treated in Ar + H_2 plasma.

Pinning

The availability of hydroxyls for bonding in a polymer molecule is of leading importance. If we consider the case of an ideal, uncoiled, linear PVA, we note that the hydroxyl unit repeats for every alternate carbon on the backbone. If we now consider such a PVA on an idealized glass surface (i.e., with a silanol for every surface silicon), we would have a matching hydroxyl ratio of PVA/glass of 1:2, or $N_m \approx 0.5$. We can thus consider, for this case, that half the silanols are "pinned".

Now consider more rigid polymers having hydroxyl groups, such

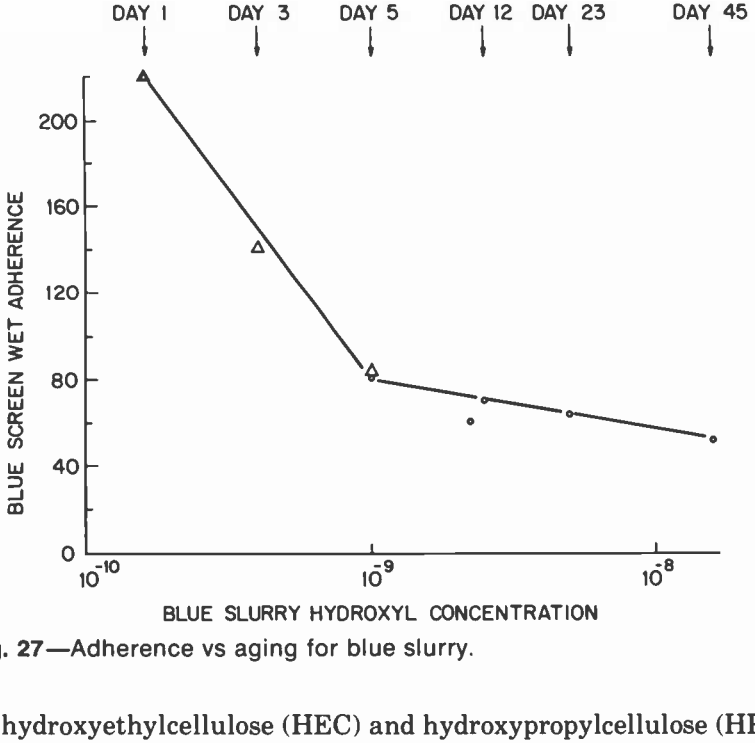


Fig. 27—Adherence vs aging for blue slurry.

as hydroxyethylcellulose (HEC) and hydroxypropylcellulose (HPC). HEC has three hydroxy groups available per anhydroglucose unit of the cellulose. Due to steric considerations, two possibilities exist for "pinning", one with $N_m \approx 0.2$ and another with $N_m \approx 0.1$. Using HEC as a precoat resulted in a screen adherence factor about six times worse than with PVA. It is interesting to compare that value to the ratio N_m^{PVA}/N_m^{HEC} which is 2.5 for one steric case and 5.0 for the second.

For the HPC case, steric hindrance and hydrogen-bonding considerations show that essentially none of the hydroxyls would be available for silanol "pinning", and our adherence experiments show no adherence when HPC is used as a precoat material.

Table 10—Plasma Treatment Effects on PVA-Precoated FPG

Plasma	Time (min)	Adherence Ratio*
Ar	1	1.00
Ar	10	0.32
Ar + H_2	1	1.00
Ar + H_2	10	1.40

* Ratio of plasma-treated sample adherence to non-treated sample.

The above data, together with our work on adherence as a function of hydroxyl content, lead to important considerations: (1) hydroxyaliphatics will yield higher "pinning" ratios than hydroxyaromatics; (2) poly(vinyl alcohol), with a "pinning" ratio of about 0.5 is a good precoat; (3) new precoats with pinning ratios greater than 0.5 should be explored.

9. Discussion

The wet adherence of phosphor screens to faceplate glass is a highly complex phenomenon involving many parameters, only a few of which have been described in this paper. The complexity is increased by the fact that many of the parameters involving adherence are not isolated but operate as interlinked systems.

Our investigations involved mainly interfaces and their effect on adherence as changes were made. Thus, we have investigated the phenomenon of etching and its effect of the glass surface. Our findings that etching with NH_4HF (or with NaOH) probably forms surface silanol groups, as well as releasing impurities from the glass, link well with the application of a PVA pre-coat whose hydroxyls can (1) react with the silanols to bond the precoat to the faceplate glass, and (2) form coordination compounds with the impurities. Other unreacted pre-coat hydroxyls react with phosphor particle PVA to bond such particles to the precoat. Interfacial chemical bonding has been discussed recently by Runge[34] for the case of glass-polybutadiene joints. The experiments we report on etchless adherence by impurity addition are of interest, particularly as adherence appeared to be a function of the solubility product of the metal sulfide. This would indicate the possibility of diffusion of the metal ion through PVA interfaces to bond with surface sulfide of the phosphor. VanOoij[35] has reported on the mechanism of rubber-to-brass adhesion which he finds due to a thin film of cuprous sulfide formed on the brass surface, with sulfur atoms being transferred to rubber molecules, thus giving interaction adhesion. Our work has indicated an optimum pre-coat film thickness for adherence and VanOoij also reports film thickness effects in his work. It is possible that our precoat optimum thickness is the limiting thickness for diffusion of metal ions. Obviously, much more work remains to be done, utilizing such techniques as ESCA, XPS, STEM, and FTIR to further characterize the surfaces and interfaces.

We have explored the effects of differences in degree of hydrolysis and degree of polymerization of the PVA and have shown the usefulness of the concept of hydrodynamic parameters to characterize

PVA. Zeta-potentials were obtained for portions of our system, and contact angles have established that successive layers of phosphor slurry see different surface conditions. Contact angle changes as a function of processing step for faceplate glass have been investigated, and the results indicated that wet adherence is dependent on contact angle and can be correlated with surface energy. The problem of crosscontamination is undoubtedly also related to surface effects of the initial and contaminating phosphors, as previous work has shown.[36] This problem has not been treated in the present work and remains to be explored in depth, particularly as it relates to surface coatings. A recent paper[37] has examined some changes in adhesion forces during picture-tube production.

The roles of double layer and of polymer inter-diffusion in adherence should be examined. At interface I_2, we have pre-coat PVA interfacing phosphor (slurry) PVA. If this is treated as the case of two identical polymers in contact, we would not expect much adherence from the doublelayer donor-acceptor aspect. Thermal motion of the molecules of the chain could lead, at the extreme case, to a mixed layer of complete mutual intersolubility. Exposing the phosphor slurry acts to cross-link the PVA, so that the phosphor PVA and the pre-coat PVA are no longer identical. However, mobility of some of the segments (although the chains are immobile translationally) can give a limited inter-diffusion, which should give a large increase in adherence. Josefowitz and Mark[38] have discussed the role of chain-segment diffusion as applied to polymer self-adherence. Interface I_2 can thus be pictured as consisting of a sensitized exposed PVA in contact with a thin PVA coating that contains ions leached from the underlying faceplate glass. Such a model for interface I_2 would provide for a limited inter-diffused region for greater adherence. The role of the exposure then would include both providing for the phosphor particle-to-particle adherence through mutual cross-linking and providing for segment inter-diffusion with the PVA pre-coat. It is of interest that Deryagin[2] gives the width of the interface diffusional zone necessary for the contact of several pairs of polymers as ranging from 6×10^{-2} mm for paraffin/SKN to 10^{-3} mm for SKB/chlorinated PVC, with an average of about 10^{-3} mm. This compares to about 10^{-4} mm for the optimum pre-coat thickness we found for PVA pre-coat on faceplate glass. In addition, since we are interested in maximizing the inter-polymer donor-acceptor pair population, it might be well to consider a double-layer system with dissimilar polymers for pre-coat and for phosphor slurry.

10. Conclusions

The adherence of phosphor screens to faceplate glass substrates is a complex phenomenon involving several interfaces and effects including the glass, the PVA, the pre-coat, and sensitization. A new method for determining adherence was evolved as part of this investigation. The effects of etching have been shown to involve silanol groups and release of impurities from the glass substrate. These, in turn, operate on the pre-coat, which has an optimum thickness and serves as an "impedance-matching" layer, binding both to the substrate and to the initial layer of phosphor particles, these interfaces being most important in adherence.

Characterization of the PVA can be carried out by determining the hydrodynamic parameters which relate to adherence. Similarly, interfacial properties can be characterized by the contact angle of the surfaces to water, resist or slurry. It has been found that phosphor slurries see different surfaces as a faceplate is screened in the shadow-mask process. The largest deviation of surface from best adherence takes place with the deposition of the second phosphor. Data on contact angle versus the screen processing step has been obtained, and adherence has been shown to be correlatable with the energy necessary to remove a screen from a surface.

An examination of the role of zeta-potentials, combined with PVA adsorption studies, indicate their possible initial importance in the adherence process, particularly with the addition of certain additives. The role of triethyleneglycol in promoting adherence was confirmed, and attributed to its effect in decreasing the sensitivity threshold.

Finally, models of etched, precoated, screened substrates were evolved.

11. Acknowledgments

The authors are grateful to the following individuals for their aid in conversations, equipment or advice: A. E. Hardy, S. A. Harper, M. R. Royce, and S. Trond of RCA Lancaster; P. N. Yocom, Y. Arie, S. Berkman, P. Datta, K. B. Kilichowski, M. Labib, A. W. Levine, H. H. Whitaker, and P. J. Zanzucchi of RCA Laboratories.

References:

[1] K. L. Mittal, Extended Abstracts, Electrochemical Society, Vol. 81-1, May 10–15, 1981, Minneapolis, Minnesota, p. 327.

[2] B. V. Deryagin, N. A. Kratova, and V. P. Smilga, *Adhesion of Solids,* Consultants Bureau (trans.), 1978.
[3] A. D. Zimon, *Adhesion of Dust and Powder,* Consult. Bur., NY, 1982.
[4] W. Fonger, private communication.
[5] M. Matsumoto and Y. Ohyanagi, *Kobunshi Kugaku* **17,** 17 (1960).
[6] H. G. Elias, *Makromol. Chem.* **54,** 78 (1962).
[7] I. Sakurada, *J. Soc. Chem. Ind.* (Japan) **47,** 137 (1944).
[8] P. J. Flory and T. G. Fox, *J. Am. Chem. Soc.* **73,** 1904 (1951).
[9] L. K. Koopal and J. Lyklema, *Faraday Disc. Chem. Soc.* **59,** 230 (1975).
[10] J. G. Pritchard, *Poly(Vinyl Alcohol),* Gordon and Breach, London, 1970.
[11] E. N. Rostovskii and L. E. De-Millo, *Zh. Prikl. Khim.* **36,** 1821 (1963).
[12] G. Natta et al., *J. Am. Chem. Soc.* **77,** 1708 (1955).
[13] C. A. Finch, *Polyvinyl Alcohol,* Wiley, London (1973), Chapt. 10.
[14] C. Y. Liang and F. G. Pearson, *J. Polym. Sci.* **25,** 303 (1959).
[15] J. F. Kenny and G. W. Willcockson, *J. Polym. Sci.* **A-1, 4,** 690 (1966).
[16] S. A. Harper, private communication.
[17] H. D. Wilcox, U.S. Patent 3,940,508, issued 2/24/76.
[18] G. Auth, private communication.
[19] M. M. Zwick, *J. Appl. Polym. Sci.* **9,** 2393 (1965).
[20] E. D. Yakhnin et al., *Makromal. Gran. Razdela Faz,* 212–217 (1971); C. A. **79,** 149805e (1973).
[21] T. F. Tadros, *ACS Symp. Colloid Dispers.,* 173 (1975).
[22] M. J. Garvey, T. F. Tadros and B. Vincent, *J. Coll. Interface Sci.* **55** (2), 440 (1976).
[23] M. L. Hair and A. M. Filbert, *Surface Chemistry of Glass,* R/D, 1969.
[24] H. L. Whitaker, private communication.
[25] See, for example, R. J. DePasquale, *Amer. Lab.,* pp. 34–39, June 1973.
[26] J. Bohemen et al., *J. Chem. Soc.* 2444 (1960).
[27] J. Relz, K. Unverferth and K. Schwetlick, *Z. Chem.* **14** (9), 370 (1974).
[28] J. T. Scholz et al., *J. Phys. Chem.* **62,** 1227 (1958).
[29] J. Frances, M. Ishaq, and M. Tatsui, *Organo Trans. Met. Chem. Jap.,* pp. 57–64 (1975).
[30] L. Grimm, K.-J. Hilke, and E. Scharrer, *J. Electro. Chem. Soc.* **130,** 1767 (1983).
[31] S. Larach and J. E. McGowan, unpublished data.
[32] W. H. Fonger, private communication.
[33] Y. Arie, private communication.
[34] M. L. Runge and P. Dreyfuss, *J. Polym. Sci.* **17,** 1067 (1979).
[35] W. J. van Ooij, *Rubber Chem. and Tech.* **51,** 52 (March-April 1978).
[36] S. Larach and J. E. McGowan, unpublished work.
[37] E. Scharrer, L, Grimm, H. Mayatepek, and W. Ritsert, *J. Electro. Chem. Soc.* **130,** 1762 (1983).
[38] D. Josefowitz and H. Mark, *India Rubber World* **106,** 33 (1942).

Operating Mechanism of the One-Piece Cathode

K. T. Chiang and J. J. Maley
RCA Video Component and Display Div., Lancaster, PA 17604

Abstract—Extensive life testing and metallurgical examinations have been conducted to study the performance characteristics and operating mechanism of the one-piece bimetal cathode. This cathode shows superior emission life and less cutoff change with life. Chromium, originally in the bottom layer of the cathode base metal, is found to diffuse into the nickel portion of the cathode cap and subsequently react with the oxide coating. A reaction in which chromium acts as an activator and chemically reduces barium oxide to free barium is proposed to account for the better emission life. The rate of chromium diffusion, rate of reaction, interface compound formation, and the effect of chromium on cutoff change during life will be discussed.

1. Introduction

Oxide cathodes for electron tubes generally consist of a nickel base substrate coated with a porous layer of alkaline earth metal oxide (BaSrCa)O. The nickel base contains a low concentration of reducing agents such as Mg, Si, Al, etc. At high temperature the reducing agents continuously diffuse into the oxide coating, reducing a small amount of barium oxide to the metal, thus providing the free barium necessary to activate the oxide emitting surface. A summary of theories advanced to explain the operating mechanism of the oxide cathode is given in Refs. [1] and [2].

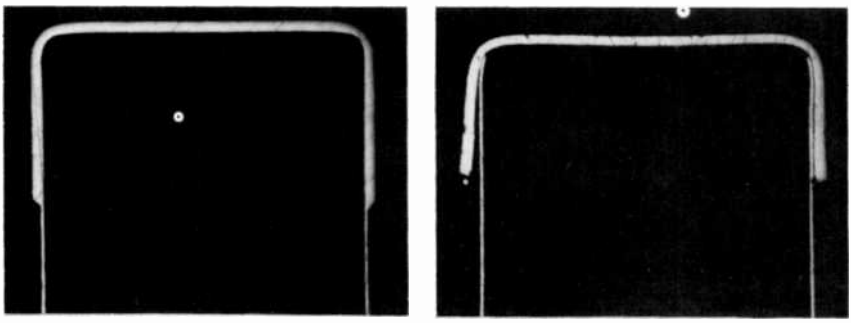
In a conventional oxide cathode, the nickel base is a cup-like structure attached to a cylindrical nichrome (Ni-20 wt.% Cr) sleeve by welding. The one-piece bimetal cathode design is an integral cathode substrate and sleeve.[3] This cathode structure may be fabricated by deep-drawing blanks from a bimetal laminate strip and

then etching away portions of the nickel alloy layer. Aside from the fabrication method, a major difference between the one-piece bimetal cathode and the conventional welded two-piece cathode is the presence of a layer of nichrome beneath the cathode nickel. This difference is illustrated by the cross-sections of the two cathodes shown in Fig. 1.

Extensive testing of the one-piece bimetal cathode was carried out to determine the emission and other electrical characteristics of the electron gun over an extended life cycle. An important parameter associated with these tests was "cutoff voltage". Given a specific negative potential applied to the first grid (G1), the cutoff voltage is defined as the positive potential required on the second grid (G2) to start drawing current of approximately 1 microampere. In tube application, the cathode should have a long emission life and a cutoff voltage that is stable over time.

The life tests were done at the normal filament voltage of 6.3 volts and at accelerated test conditions of 6.9, 7.5, and 8.1 volts. The major conclusions from these tests were that (1) the emission life of the one-piece bimetal cathode is superior to that of the conventional two-piece cathode and (2) the cutoff change with life is less for the one-piece bimetal cathode than for the two-piece cathode. Results of the typical life tests for emission and cutoff voltage are presented in Figs. 2 and 3, respectively.

To explain these test results, we initiated a program to study the operating mechanism of the one-piece bimetal cathode. An under-



a b

Fig. 1—(a) Cross-section of the one-piece cathode formed by removal of nickel from the sleeve portion of the bimetal; the cathode cap has 2 mils of Ni-4 wt% W on top of 1 mil nichrome. (b) Cross-section of standard KS904 cathode; the 2-mil Ni-4 wt% W cap is welded to the nichrome sleeve. (Enlarged approximately 20×.)

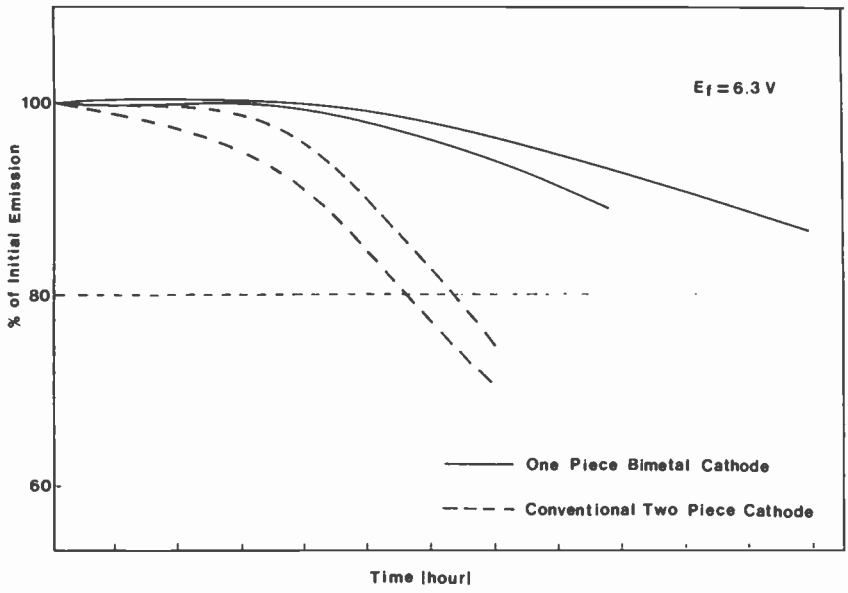


Fig. 2—Cathode emission versus time.

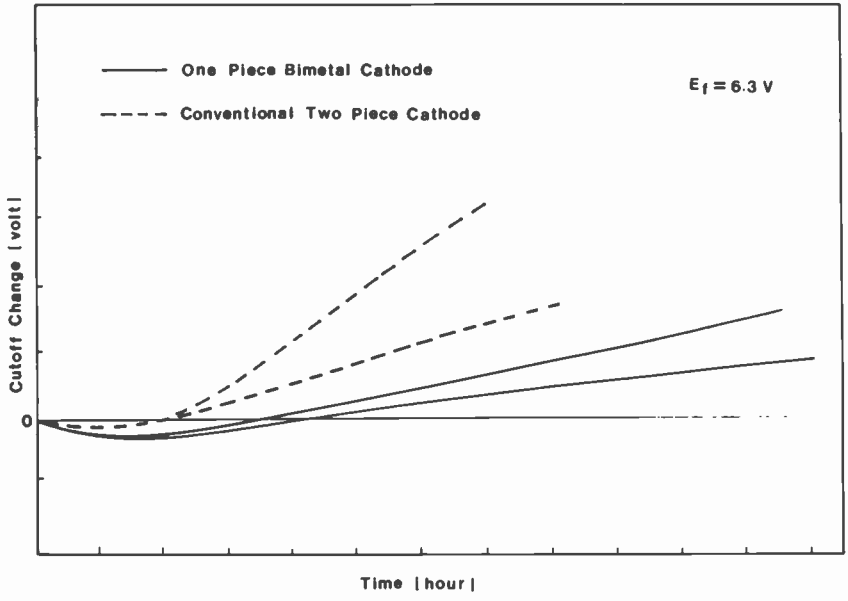


Fig. 3—Cutoff voltage change versus time.

standing of the operating mechanism is essential to the manufacture, processing, and use of this cathode, as well as for future cathode development. A survey of scientific literature on oxide-coated cathodes[4,5] revealed that chromium above 0.003 wt.% has never been intentionally added or present in cathode nickel. Special attention was therefore paid to the study of the migration of chromium and its effect on cathode performance. Specifically, it was intended to answer the following questions concerning the presence of chromium in the one-piece cathode.

(1) Earlier researchers[6] have reported the formation of a high-resistance interface layer on cathodes made from chromium-plated nickel. Does a similar reaction occur on the one-piece cathode?
(2) How long does it take for chromium to diffuse through the nickel portion of the cathode cap at normal operating voltage of 6.3 V E_f and at accelerated test conditions?
(3) After reaching the oxide-coating/metal interface, how does chromium react with the oxide coating? Does the reaction compound form an interface barrier?
(4) What is the role of chromium in improving the emission and cutoff characteristics of the cathode during life?

All of these aspects were carefully examined during the course of this study. As a result, an operating mechanism that extended over all stages of the cathode life was defined.

2. Experimental Procedures

The experimental methods used for analysis involved x-ray diffraction, optical metallography, x-ray photoelectron spectroscopy (XPS), scanning electron microscopy, and electron microprobe analysis.

Most long-term-life-test samples had at least 15,000 hours of life. After life test, the cathodes were visually inspected to determine the extent of coating discoloration. Cross-sections of the cathodes were then prepared for further analysis using optical-microscope and scanning-electron-microscope techniques. For selected samples, parts of the oxide coating were scraped from the cathode and analyzed by x-ray diffraction. This technique offered information on the crystal structure of the reaction product and allowed new phases to be identified.

To study chromium diffusion in the base metal, microprobe scans were made on cross-sectional samples along a line perpendicular to the cathode surface. A series of samples with different life-test con-

ditions and operating times were examined. The tube operating conditions included filament voltages (E_f) of 6.3, 6.9, 7.5, and 8.1 V. The corresponding cathode temperature, in terms of optical brightness, were approximately 1075, 1120, 1150, and 1200°K. For operating filament voltages of 6.3 and 6.9 V, the cathodes were examined after interrupting the life tests at 1, 3, 5, 11, and 15 weeks. At filament voltages of 7.5 and 8.1 V, the life tests were interrupted at 1, 2, 3, 5 and 7 weeks.

Cathodes operated for even longer times were chosen for the study of the migration of chromium into the oxide coating. Energy dispersive analysis of x-rays (EDAX) was used to identify the elements present in the coating. Scanning electron micrographs and x-ray dot maps were taken to show the distribution of chromium in the base metal as well as in the oxide coating. In this analysis, the detector was tuned to respond to the x-ray wavelength characteristic of chromium. The density of the white dots therefore provided information on the abundance of chromium in the examined areas. To study the interface compound between the oxide coating and the base metal, similar x-ray maps of barium, tungsten, oxygen and nickel were also obtained.

XPS analysis was carried out to study the cathode surface before spraying the alkaline-earth carbonate coating. This surface-sensitive technique offered chemical-element and chemical-bonding information within a 100 Å range of the sample surface.

3. Initial Reaction at the Metal/Coating Interface

To study the surface reaction between the oxide coating and the base metal, three sets of cathode samples were prepared for XPS analysis:

(1) One-piece bimetal cathode before H_2 firing
One-piece bimetal cathode after H_2 firing
(2) Nichrome cathode before H_2 firing
Nichrome cathode after H_2 firing
(3) Standard KS904 cathode before H_2 firing
Standard KS904 cathode after H_2 firing

The nichrome cathode was made by removing all the nickel from the one-piece bimetal cathode. XPS analysis showed the following results:[7]

(1) Approximately 10 nm of nickel oxide was observed on the unfired nickel surface of the one-piece bimetal cathode, composed of NiO and a nickel suboxide Ni_2O_3 (or perhaps a partial hy-

droxide NiOOH). After H_2 firing, the oxide was somewhat thinner and richer in NiO. The oxide on the standard cathode closely resembled the one-piece cathode.

(2) The surface of the unfired nichrome cathode was chromium enriched and contained a mixture of Cr_2O_3 and nickel oxides. No nickel-oxygen bonding was present after firing, but the Cr_2O_3 was unaffected.

The results are as expected from a thermodynamic perspective. The atmosphere of the hydrogen firing furnace is slightly reducing for nickel oxide on the cathode caps of the standard and the one-piece bimetal cathode. For the nichrome cathode, however, Cr_2O_3 is present on the cathode surface and is stable in the H_2 firing atmosphere. The Cr_2O_3 enriched layer continues to grow during subsequent cathode processing and tube processing. As a result, after the cathode activation process at a cathode temperature of about 1250°K, the interface reaction compound on the nichrome cathode surface is of the form $BaO\text{-}Cr_2O_3$. This interface layer is similar to that in chromium plated nickel. The higher resistance of this compound[6] may cause difficulty in drawing current through the coating. In addition, the coefficient of thermal expansion of this layer may be different from the cathode substrate body causing peeling of the oxide coating. Both these effects were observed for nichrome cathodes at 8.1V E_f life test. The nichrome cathodes had early emission failures at 8.1V E_f life test and many of the cathodes had poor coating adherence. For the standard cathode and the one-piece bimetal cathode, the interface reaction between the oxide coating and the substrate is of the form BaO-NiO. This reaction forms a strong bond between the oxide coating and base metal and is responsible for the good adherence of the cathode coating.

This analysis shows that the detrimental effect of high-resistance chromium compound formation was avoided due to the presence of a 2-mil layer of cathode nickel.

4. Chromium Diffusion in the Base Metal

4.1 During Tube Processing

Fig. 4a shows an optical micrograph of the one-piece cathode before tube processing. The two-layer structure is clearly visible. The 2-mil cathode nickel (with 4 wt.% W and trace amounts of Mg and Si as reducing impurities) was hot bonded to the 1-mil nichrome alloy (Ni-20 wt.% Cr). Microprobe traces of Ni and Cr (Fig. 4b) across the cap indicated chromium was confined to the bottom 1-mil layer, and

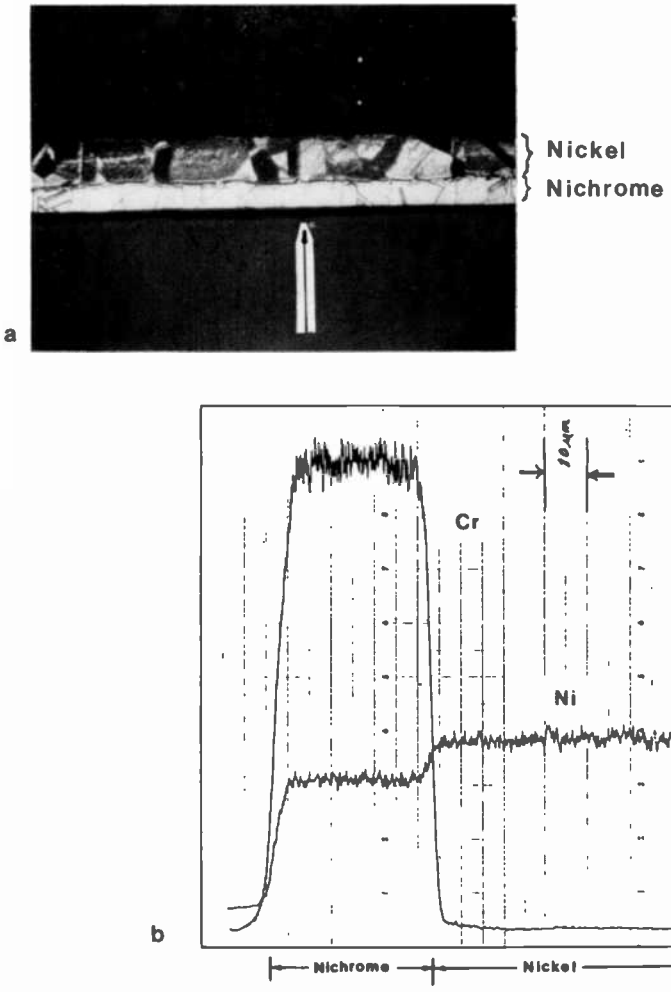


Fig. 4—(a) Cross-section of the one-piece cathode cap before tube processing (approx. 130×). (b) Ni and Cr concentration profiles along the line of traverse indicated by the arrow in the photo.

the top layer contained higher nickel concentration. This result is as expected, since the initial cathode nickel contained 96 wt.% nickel as compared to about 80 wt.% in the nichrome layer. After tube processing, a significant amount of chromium had already diffused into the nickel portion of the cap. At the same time, some nickel diffused back to the nichrome portion. The interdiffusion zone is seen in the cross-section (Fig. 5a) as the slightly darker area

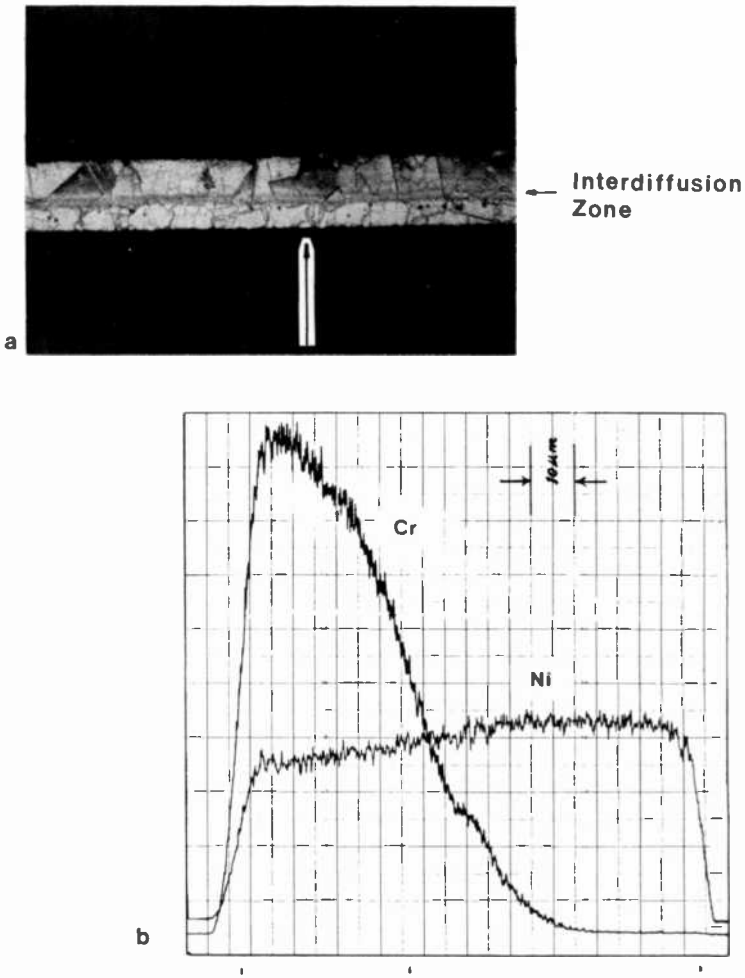


Fig. 5—(a) Cross-section of the one-piece cathode cap after tube processing (approx. 130×). (b) Ni and Cr concentration profiles along the line of traverse indicated by the arrow in the photo.

around the original nickel/nichrome interface. The interdiffusion of nickel and chromium is clearly evident from the concentration profiles of the two elements as shown in Fig. 5b.

4.2 During Life

Chromium profiles across the cathode cap after different operating times at filament voltages of 6.3, 6.9, 7.5, and 8.1 volts are presented

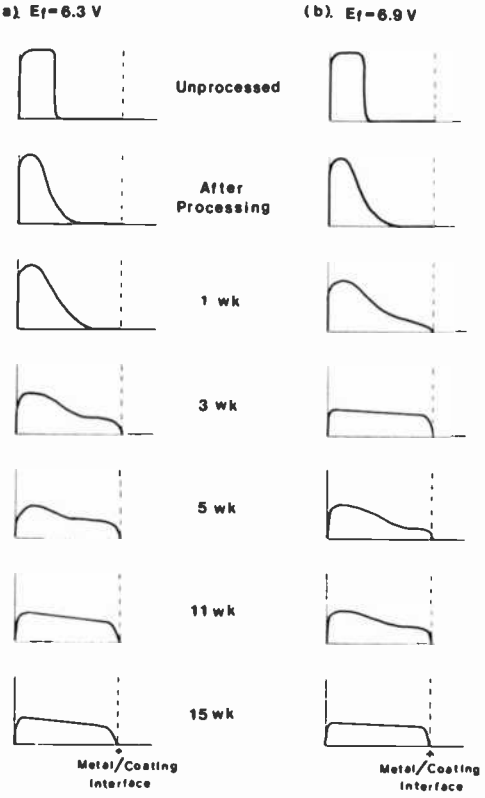


Fig. 6—Cr concentration profiles across the one-piece cathode cap after different times on life at filament voltages of 6.3 and 6.9 volts.

in Figs. 6(a), (b), (c), and (d). Different samples were examined for different operating voltages and operating times. It is noted that in all cases, chromium diffusion progresses at a rapid rate. As shown in Fig. 6(a), at a filament voltage of 6.3 V, it takes about 3 weeks for chromium to diffuse through the nickel portion of the cap and to reach the metal/oxide coating interface. At an operating filament voltage of 8.1 V, it takes only one week for chromium to diffuse to the interface (Fig. 6(d)). Similarly, only one to two weeks are needed for chromium to reach the metal/oxide coating interface at filament voltages of 6.9 and 7.5 V (Figs. 6(b) and 6(c)). This rapid diffusion rate can be explained by diffusion through easy diffusion paths, such as grain boundaries that are present in both the nickel and nichrome portion of the cap. The grain structure in the cathode cap is shown in Fig. 7a. The grain structure in the nichrome alloy is

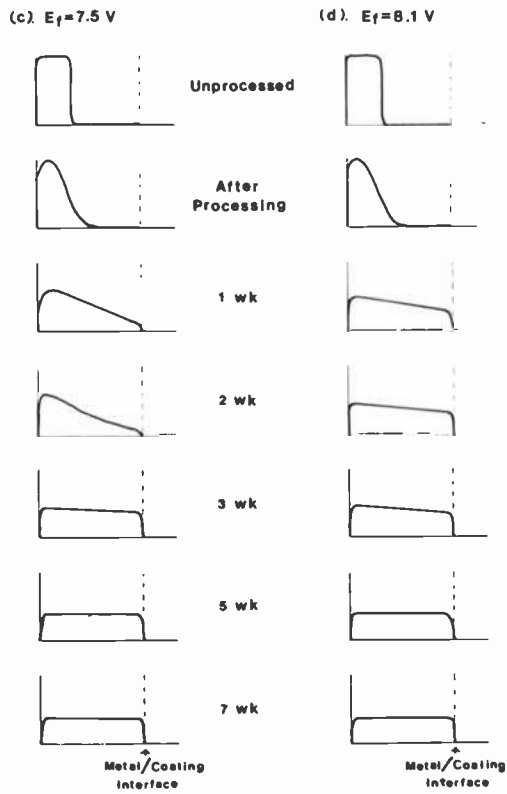


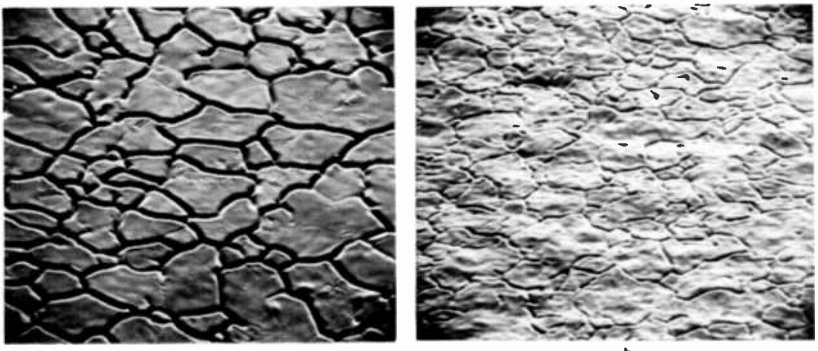
Fig. 6—(Continued)—Cr concentration profiles for filament voltages of 7.5 and 8.1 volts and for different periods of time on life than in Fig. 6(a) and (b).

shown in Fig. 7b. The rapid diffusion along these grain boundaries, followed by volume diffusion within the grains, allows chromium to diffuse into the nickel lattice during early stages of the cathode operation.

5. Reaction Between Chromium and Oxide Coating

5.1 Migration of Chromium

Two one-piece bimetal cathodes after long term life tests at 6.3 and 5.7V E_f and one similar cathode after accelerated life test at 8.1V E_f were selected for this study. All three cathodes out-lived their two-piece cathode counterparts. The operating conditions, time on



a b

Fig. 7—Surface micrograph of the one-piece cathode showing grain boundaries within the base metal (approx. 100×): (a) grain structure of cathode cap and (b) grain structure of nichrome alloy.

life, and cathode emission levels when the tubes were opened, are as follows:

E_f = 6.3V, 110 weeks, greater than 80% of initial emission
E_f = 8.1V, 10 weeks, approximately 80% of initial emission
E_f = 5.7V, 110 weeks, greater than 80% of initial emission

The heating cycle of the life test is 2-3/4 hours on and 1/4 hour off. One hundred and ten weeks of life test corresponds to about 17,000 hours of cathode working life.

Fig. 8a shows the cross-section of the 6.3V E_f cathode. EDAX of the outer part of the coating shows Ba, Sr, and Ca, which are the constituents of the oxide coating (Fig. 8b). The inner part of the coating contains Cr in addition to Ba, Sr, and Ca (Fig. 8c). The migration of chromium into the oxide coating is clearly indicated. X-ray powder diffraction of the coating scraped from the 6.3V E_f sample and from 8.1V E_f samples showed the chromium-containing phase was in the form of $Ba_3(CrO_4)_2$.[8] This phase is yellow and can be visually detected in cathodes after accelerated life test.

To compare the extent of chromium migration into the oxide coating at different operating temperatures, chromium x-ray maps were taken of the three samples. Fig. 9a shows that after 110 weeks at 6.3V E_f, a small amount of chromium was present in the inner part of the coating. For the cathodes after 10 weeks at 8.1V E_f, massive migration of chromium into the coating was observed (Fig. 9b). In some areas, migration through the oxide was complete. The chromium concentration in the oxide coating was estimated to be between 10–50 times higher than in the chromium-containing region of the 6.3V E_f cathode. In contrast, no detectable chromium

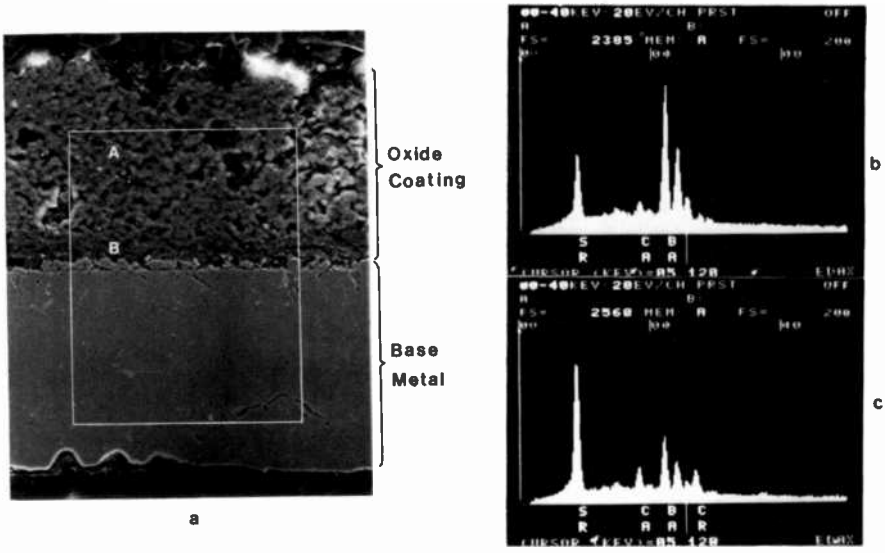


Fig. 8—(a) Scanning electron micrograph of the one-piece cathode after 110 weeks at E_f = 6.3 V (approx. 400×); (b) shows EDAX of outer part of oxide coating (point A) and (c) EDAX of inner part of oxide coating (point B).

was found in the oxide coating of the 5.7V E_f cathode (Fig. 9c). Although Cr may be present in the oxide in a small amount, it is below the detection limit of this technique. Some background x-rays are seen in the coating part of the micrograph.

The aforementioned results show that the rate of chromium migration is highly temperature dependent in the temperature range of the life tests. Other significant results from these data are:

(1) In all three cases, chromium is uniformly distributed in the base metal including the portion originally containing only cathode nickel. (See Figs. 9a, b, c.) These data confirm the results of microprobe trace study of diffusion of chromium in the base metal.
(2) In all three cases, there was no accumulation of chromium in the oxide-coating/metal interfaces. (See Figs. 9a, b, c.)

If the reaction between chromium and the oxide coating is controlled by solid-state diffusion of chromium in the base metal, a high concentration of chromium in the interface is expected. The observation that chromium works its way up through the highly porous cathode coating instead of accumulating in an interface layer

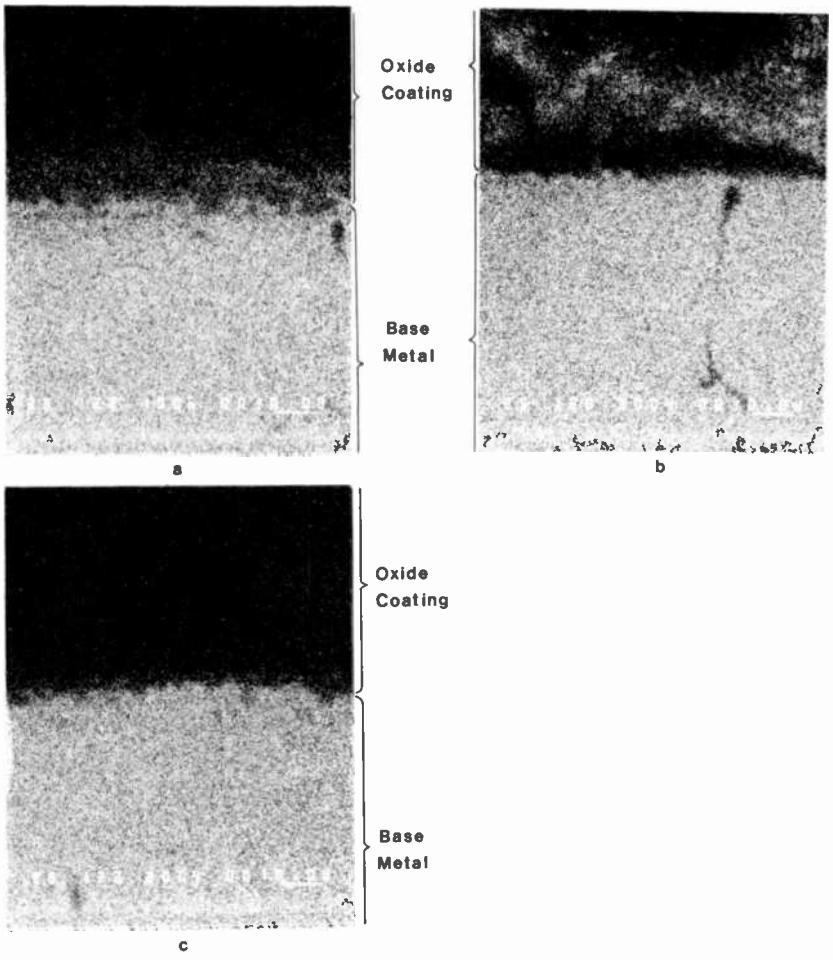


Fig. 9—Chromium Kα x-ray maps of the one-piece cathode: (a) after 110 weeks at E_f = 6.3 V (area outlined in Fig. 8a), (b) after 10 weeks at E_f = 8.1 V, and (c) after 110 weeks at E_f = 5.7 V.

strongly suggests that the migration is through a vapor transport process.

The vapor pressure of chromium at various temperatures can be calculated from Ref. [9]. At filament voltages of 5.7, 6.3, and 8.1V, the vapor pressures of chromium are 4.08×10^{-11}, 8.05×10^{-11}, and 4.02×10^{-9} atm, respectively. The vapor transport rate of chromium was increased by a factor of 2 when the filament voltage was increased from 5.7V to 6.3V, and by a factor of 50 when the

filament voltage was increased from 6.3V to 8.1V. These data show that the vapor transport mechanism is consistent with the observed temperature-dependent nature of the chromium migration.

5.2 Interface Compound

To complete the description of the coating/metal interface reaction for the one-piece cathode, nickel, tungsten, barium, and oxygen x-ray maps were taken of the interface region of the 6.3V E_f cathode. The region examined is the blocked area shown in Fig. 8a. Nickel stays in the base metal (Fig. 10a). An accumulation of tungsten in the coating/metal interface is clearly seen (Fig. 10b). This layer is also enriched with barium and contains oxygen (Figs. 10c and 10d), suggesting that the interface compound is barium tungstate (possibly $3BaO \cdot WO_3$). The formation of the barium tungstate interface layer is expected and observed for all KS904 cathodes using Ni-4 wt.% W cathode nickel.[10] Some penetration of barium and oxygen into the metal substrate is also noted in Figs. 10c and 10d.

This analysis shows the interface compound of the one-piece bimetal cathode to be similar to the standard KS904 cathode. Chromium does migrate into the coating and reacts with the oxide to form a compound of $Ba_3(CrO_4)_2$. The reaction is controlled by vapor transport of chromium and does not concentrate in an interface layer. Cathodes with this compound in the oxide coating have very long emission life. This is further evidence that the chromium migration does not cause excessive bulk or interface resistance or loss of coating adherence.

6. Reaction Mechanism

6.1 Role of Chromium in Reducing Cutoff Change

Cutoff voltage is an electrical characteristic of the electron gun that is sensitive to cathode-to-grid (G1) spacing.[11] An increase in cutoff voltage during life usually represents an increase in cathode-G1 spacing. Contributors to this increase are sintering of the cathode coating during life, thinning of the coating due to evaporation of coating material, and the thermal deformation of cathode cap and sleeve. Comparing the one-piece bimetal cathode and the standard two-piece cathode, the effects of coating sintering and evaporation are the same. The initial slight decrease in cutoff voltage of the one-piece cathode in life testing may be due to the elimination of cathode cap-sleeve welding and related to the interdiffusion of nickel and chromium. Details of this process are not yet completely understood.

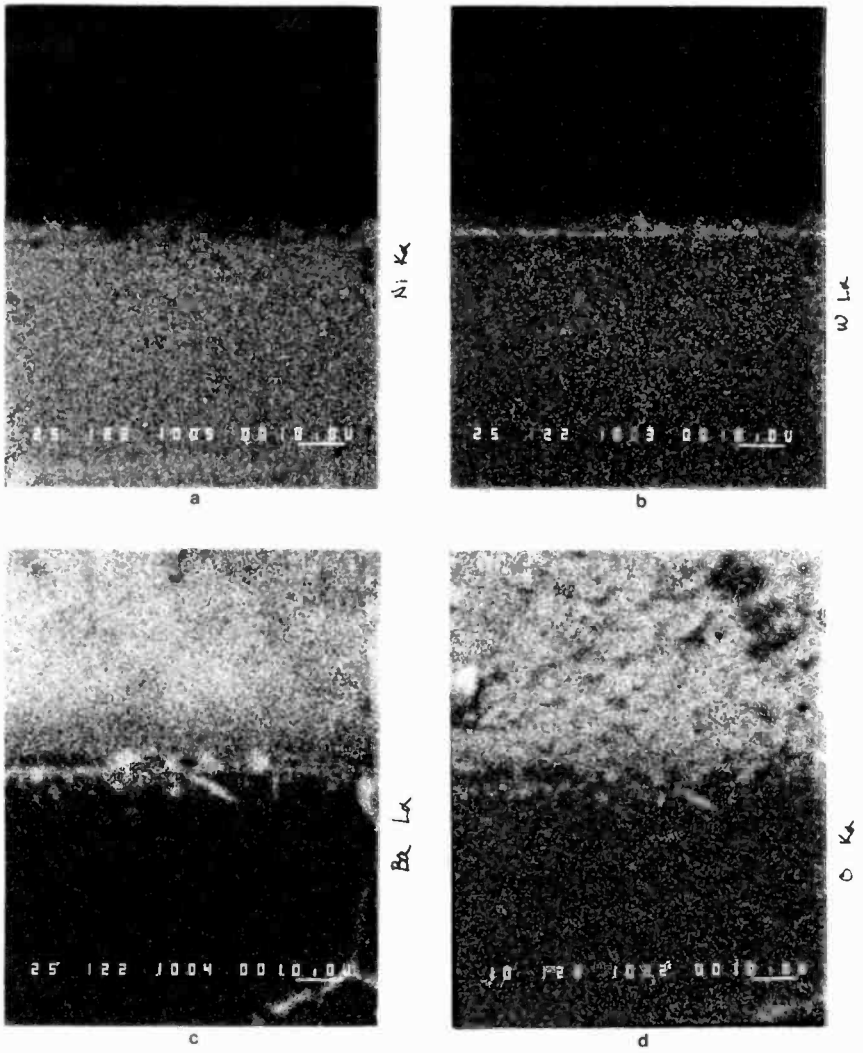


Fig. 10—Distribution of nickel, tungsten, barium, and oxygen in the area outlined in Fig. 8a: (a) Ni Kα x-ray map, (b) W Lα x-ray map, (c) Ba Lα x-ray map, and (d) O Kα x-ray map.

After several weeks of life test, nickel and chromium are uniformly distributed throughout the cathode cap by diffusion. The presence of chromium in the nickel alloy strengthens the nickel lattice and results in smaller cap deformation during life. Consequently, the net cutoff increase in the same operating time period

is reduced. The solid-solution strengthening effect of chromium on nickel base alloys at high temperature is well documented in the literature.[12,13]

6.2 Role of Chromium in Increasing Cathode Life

Maintenance of emission levels is dependent upon a delicate balance of the amount of free alkaline-earth metal (Ba or Sr). A continuous production of free Ba is required, because a constant evaporation of the metal from the oxide surface takes place during tube operation. The production of free Ba is determined by the transport and availability of the activators. As the operating time of a cathode increases to thousands of hours, the supply of activators (Mg or Si) can be gradually depleted in two ways.

First, evaporation and activation reaction at the coating/metal interface gradually deplete the core metal of its activators.[14] Consequently, the flux of activators to the interface is reduced. Second, BaO dissociates into Ba and O. Oxygen and part of the barium diffuse back into the nickel cap.[15] Mg and Si may be oxidized inside the base metal through the reactions:

$$Mg + O \rightarrow MgO$$

$$Si + 2O \rightarrow SiO_2$$

Barium may precipitate along these oxides at grain boundaries. Subsequent diffusion of free activators is partly inhibited by these internal oxides.

For the one-piece bimetal cathode, chromium diffuses into the nickel portion of the cathode cap during the first several weeks of the tube operation. After diffusing to the coating/metal interface, chromium migrates via vapor transport to react with the oxide coating. A reaction such as

$$8BaO + 2Cr \rightarrow Ba_3(CrO_4)_2 + 5Ba$$

occurs in the coating. In this reaction, chromium acts as an activator to chemically reduce barium oxide and thus provide an excess of free metallic barium. The life of the cathode is prolonged by this reaction when original activators are depleted by the two aforementioned processes.

Although there is not much difference in the time needed for chromium to diffuse to the oxide/metal interface at different operating temperatures (about 1 week at 8.1V E_f, 3 weeks at 6.3V E_f), the migration rate of chromium into the oxide coating changes greatly with different operating temperatures. At the normal op-

erating temperature of 6.3V E_f, chromium is metered out slowly and the cathode activity is maintained over a long period of time.

7. Summary and Conclusions

The one-piece bimetal cathode is a system that shows superior emission life and less cutoff voltage change with time than a conventional cathode. The major structural difference between the one-piece bimetal cathode and the standard two-piece cathode is the presence of a 1-mil nichrome layer beneath the cathode nickel.

Initial activation of the one-piece cathode is accomplished by the activators (Mg and Si) present in the cathode nickel, as in the case of the standard cathode. The interface reaction between the oxide coating and the substrate is of the form BaO-NiO. During tube processing and the initial stage of cathode life, chromium diffuses into the nickel portion of the one-piece cathode cap. The process involves grain boundary diffusion and progresses at a rapid rate. Only one to three weeks are needed for chromium to reach the oxide coating/metal interface. Chromium subsequently migrates via vapor transport to react with the oxide coating. In this reaction chromium acts as an activator and prolongs the cathode emission life.

The decrease in cutoff voltage change with time for the one-piece cathode is related to the elimination of cathode cap-sleeve welding and to solid-solution strengthening of the nickel lattice by chromium.

Acknowledgments

The authors gratefully acknowledge the contributions of J. C. Bridgeman, L. J. Royer, R. H. Woods, K. M. Finlayson, J. M. Wolfe and Dr. R. W. Heiden of the RCA VC&DD Analytical Laboratory and S. J. Kilpatrick of RCA Laboratories' Analytical Laboratory.

References:

[1] G. Hermann and S. Wagener, *The Oxide-Coated Cathode*, Volumes 1 and 2 (1951), Chapman and Hall, London.
[2] P. Zalm, "Thermionic Cathodes," in *Advances in Electronics and Electron Physics*, **25**, pp. 211–275 (1968).
[3] J. C. Turnbull, RCA Technical Note 1159 (1976).
[4] C. H. Meltzer and E. G. Widell, "Nickel Base Metal for the Oxide Cathode," in *Electron Tube Design*, pp. 64–76 (1962).
[5] W. H. Kohl, "Cathodes and Heaters" in *Handbook of Materials and Techniques for Vacuum Devices*, Chapter 16 (1967).
[6] A. Fineman and A. Eisenstein, "Studies of the Interface of Oxide Coated Cathodes," *J. Appl. Phys.*, **17**, p. 663 (1946).

[7] S. J. Kilpatrick, Private Communication.
[8] JCPDS Powder Diffraction File (1979).
[9] *JANAF Thermochemical Tables*, 1975 Supplement, Reprint No. 60 from *J. Physical and Chemical Reference Data*, **4**, (1975).
[10] E. J. Hannig, "The Interface and Other Oxide-Cathode Impedances," in *Electron Tube Design*, pp. 77–89 (1962).
[11] H. Moss, in *Narrow Angle Electron Guns and Cathode Ray Tubes, Advances in Electronics and Electron Physics Supplement 3* (1968).
[12] K. Jackson and R. O. Jenkins, "Nickel in Electronics," *Metallurgia*, **47**, p. 277 (1953).
[13] R. M. Brick, R. B. Gordon, and A. Phillips, "Metals for High Temperature Service," in *Structure and Properties of Alloys*, Chapter 14 (1965).
[14] E. S. Rittner, "A Theoretical Study of the Chemistry of the Oxide Cathode," *Philips Res. Repts.*, **8**, p. 184 (1953).
[15] A. C. Aten, J. C. Duran, J. J. Guerts, and A. J. Griffioen, "Chemical Transport in Oxide Cathodes," *Philips Res. Repts.*, **26**, p. 519 (1971).

A Contour Deformation Model of Capacitance VideoDisc Signal Pickup

P. D. Southgate
RCA Laboratories, Princeton, NJ 08540

Abstract—A method of calculating the videodisc pickup signal is described that includes the elastic deformation of the disc by the stylus. The stylus shoe is idealized to a flat, rectangular shape and small vertical displacements are assumed. Disc-stylus capacitance is calculated using a two-dimensional geometry and an approximation to the field configuration that allows fast computation. Typical results are described for the wavelength and tracking-force dependence of pickup. There is agreement between most general features of calculated and measured pickup. The mechanism of soundbeat is investigated and shown, within the terms of this model, to depend primarily on the elastic deformation of the signal contour. The calculated magnitude and the wavelength dependence of soundbeat agrees well with measurement; for a "flat-top" signal, drop of soundbeat and change of wavelength dependence of the type observed are predicted. Calculations usually deviate from measurements only by a magnitude similar to the mutual deviation of measurements made under varying conditions.

Introduction

The successfully optimized combination of stylus configuration and signal depth used in the capacitance pickup VideoDisc system has been developed by exhaustive experimental trial in conjunction with idealized models of electrode-disc capacitance. These models, although providing substantial insight into the pickup mechanism, are not complete, since they do not fully take into account the elastic deformation of the disc surface by the stylus. Since, as will be seen, deformation has a major effect in modifying the signal contour, particularly with regard to second-order effects, such as harmonic gen-

eration and soundbeat, it is of considerable interest to include this deformation. In this paper, a model is developed that represents the deformation, with some approximations that are geometrical in nature and are believed to be small. The model is simplified by representing the disc surface as a perfect conductor with a thin dielectric layer on top; gradations and fluctuation noise in the conductive layer are not included. It is also assumed that the stylus tip is surrounded by the lubricating oil so that signal-frequency fluctuations in the oil surface contour do not affect the electrode-disc capacitance.

Some general features of the signal pickup predicted by the model will be described and compared with measurements made by a number of workers in RCA Laboratories. The correspondence with measurement is usually reasonably good, within the variability of the measurements. Small systematic differences may be caused by the geometric approximations made in this analysis; a greater difference probably arises from an inadequate accounting of the oil film behavior, both in its viscoelastic properties and its varying dielectric effect. A full treatment of this would be very difficult; the model is therefore presented as a demonstration of how closely stylus performance can be predicted with only an elementary inclusion of these effects.

Other models, using different approximations, have been described in the past. The basis of these has been indicated by Clemens;[1] the analysis has been extended considerably in unpublished work by J. J. Gibson, J. R. Matey, and R. W. Nosker.

Elastic Contour Deformation

Computation of the deformation due to the stylus is reduced to manageable proportions by using an idealized stylus form. Fig. 1 compares this form with the real stylus. The grooves are flattened and the contact area of the shoe has a straight rather than an acute leading edge. The shoe shape transverse to the groove direction is taken to be slightly rounded so that the pressure under the shoe does not depend on the distance from the groove center. Calculations of vertical surface-contour deformation are made at points along the groove center; it can be shown that if the location is moved toward the side of the shoe, the deformation will not change much until it is close to the side.

The elastic modulus used is the low-strain time-independent value; yield, relaxation effects, and inertial effects are ignored. Although the total stress near the stylus contact at the tips of the

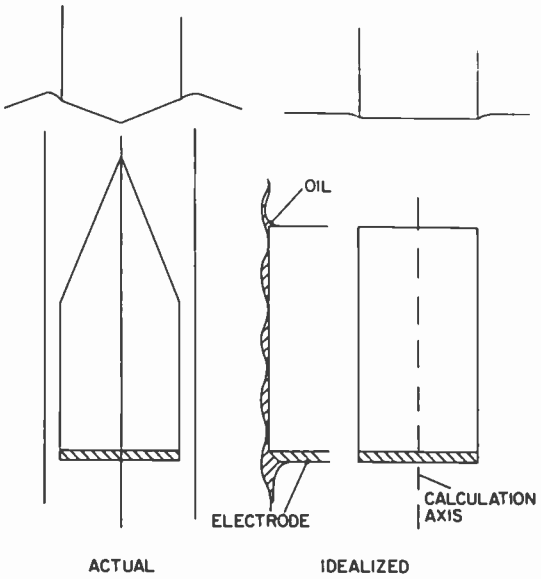


Fig. 1—Idealized form of stylus, used in calculations, compared with real stylus. Electrode and signal depth is approximately to scale; distribution of oil is assumed.

signal elements can be large, the smaller shear component, which would lead to yield, usually does not exceed the yield stress. In addition, it is transient, and the short-term yield stress in disc material has been shown to be much higher than the static value. The material damping has been shown to be low at frequencies of the order of 1 MHz, justifying the neglect of relaxation effects. Inertial effects are not important, because the speed of sound is about 400 times the speed of the disc, so that the elastic deformation is always near equilibrium.

In addition to its possible dielectric effects, the lubricating oil can modify the pressure distribution under the shoe if it is sufficiently viscous. The bulk viscosity is of the order of 0.2 poise at low shear rates and probably considerably less at the high rates obtaining between the shoe and disc. Simple calculations show that for normal video signals, where the channel space between the shoe and the disc is 800 Å or more, the normal stylus pressure will cause the oil to flow laterally outward in a time short compared with the transit time of the disc under the shoe. The pressure profile near the electrode will therefore be little modified by the oil. For signals of shallow modulation depth, however, oil flow may be slower; for a smooth groove and oil of viscosity 0.1 poise, an initial thickness of

400 Å will be reduced to 250 Å only after the oil has traversed the whole way under a 4μm long shoe. It is likely therefore that the oil under these conditions will cause the pressure under the leading part of a shoe, which is essentially parallel to the disc, to be a little greater than that under the following part. The oil may also contribute an essentially elastic thin surface layer; this layer will be included in the surface dielectric of the disc material. Errors introduced by the above approximations are believed to be small but are difficult to assess; it would require a comparison with exact calculation for some configurations, a formidable task which has not been attempted.

A point force F perpendicular to a horizontal semi-infinite elastic solid surface produces a vertical deflection at distance r, if F is small, of

$$z = F(1 - \nu^2)/\pi E r, \qquad [1]$$

where E is Young's modulus and ν Poisson's ratio. If the stylus shoe is divided into strip elements as shown in Fig. 2, then the vertical deflection at a point on the groove center due to one of these elements may be obtained by integrating the deflection of Eq. [1] over the element, which is assumed to apply a constant vertical pressure to the disc surface. It is convenient to use two analytic approxi-

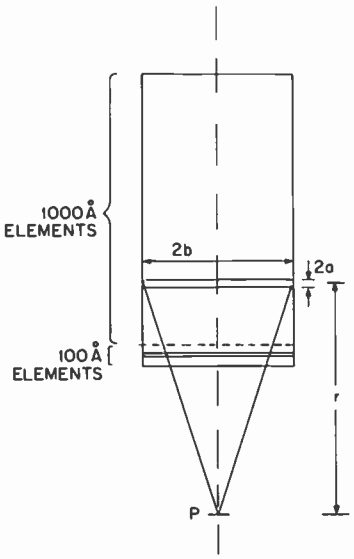


Fig. 2—Shoe-area element dimensions.

mations to this integral. One is for a close element of width $2a$ comparable to its distance r from the point considered.

$$z = \frac{F}{b\pi}\frac{(1-\nu^2)}{E}\left(\frac{1}{2a}[(r+a)\ln(r+a) - (r-a)\ln(r-a)] - 1 - \ln 2b\right); \quad [2]$$

the other approximation is for more distant elements,

$$z = \frac{F}{b\pi}\frac{(1-\nu^2)}{E}\frac{1}{2}\ln\frac{(1+r^2/b^2)-1}{(1+r^2/b^2)+1}. \quad [3]$$

Here F is the total force on the element and a and b are its half width and length. The b/a ratios used in these calculations (20 or greater) give a sufficiently large region of overlap between these two approximations.

If the number of elemental spaces between the force and the deflection point is $k = r/2a$, then the interaction coefficient is defined as $I(k) = z/F$, given by the appropriate Eq. [2] or [3]. The total displacement at the location of element j along the axis due to a set of forces $F(i)$ at location i will then be

$$z(j) = \sum_i F(i)\, I(i-j). \quad [4]$$

Incorporation of a disc surface signal modulation into these calculations proceeds as in Fig. 3. The upper diagram shows the signal surface contour before stylus pressure is applied, the mean plane of the disc being indicated by the broken line. When the pressure is applied, the disc mean plane is depressed by the sum of the stylus depth below the initial mean plane and the signal contour modulation, if the stylus is in contact with the disc. The situation is shown in the lower diagram. In these calculations, the elastic strain in those parts of the disc between the surface and the mean plane is ignored. Region A (Fig. 3) is assumed to be stiff in the vertical direction (so that $d_1 = d_2$) but not to contribute otherwise to the disc stiffness; region B is taken to contribute to disc stiffness although it is in fact not present. Fig. 3 suggests a greater error from this cause than is the case, since the vertical scale is exaggerated; the error can be shown to be small and comparable to that produced by the implicit assumption of Eq. [1] that the vertical displacements are small compared with the horizontal scale.

The procedure followed to calculate surface contour deformation, then, is first to consider only those elements in contact with the shoe. The vertical displacements of the disc surface at the locations

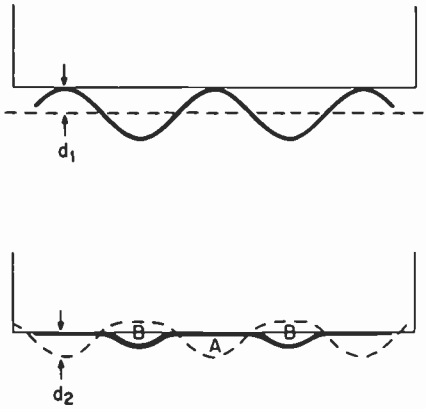


Fig. 3—Elastic deformation of mid-plane of disc surface showing regions of material in which the compression is ignored.

will be the sum of the stylus depth below the mean disc plane and of the modulation. Thus, in the set of Eqs. [4], if the range of i and j is restricted to only the contact points, all values of $z(j)$ are known, so that this forms a complete and solvable set of equations. The set may then be inverted to give the $F(i)$ values.

This procedure is followed iteratively. At first, a convenient set of points under the stylus but near the signal peaks is chosen. Eqs. [4] are solved and the derived $F(i)$ values are used to obtain the surface contour at all points under the stylus. Two physically unrealistic situations may then be found: the disc surface may be above the shoe or the force may be negative (so that the surface is pulled up toward the shoe). Points of the first kind are added to the original set and points of the second kind are removed, forming a new set of final contact points. Further iterations are then made until all forces are positive and all disc surface locations are at or below the shoe surface. Given this final set of forces, the desired calculation of the deflection of all points on the disc surface can be made.

It is only close to the electrode that a small element size is needed to give accuracy in the final capacitance calculations. In the calculations to be described, the elements within 2000 Å of the electrode-stylus interface are 100 Å wide, and elsewhere they are 1000 Å, as indicated in Fig. 2. Interaction coefficients are scaled to accommodate this change of size.

To produce a varying stylus tracking force, the depth of the shoe below the average disc surface is varied. It is a feature of these computations that the tracking force cannot be specified in advance;

reasonable estimates must be made and the final results for the desired tracking force found by interpolation. The curvature of the stylus shoe can vary depending on its playing-wear history. An assessment of the stylus shoe profile can be made from scanning electron microscope (SEM) observations. When this is known for the particular shoe used, or for one that has had a similar history, an analytic form can be fitted to the observed profile and this form can be used in the deformation calculations. Alternatively, a "standard" shoe profile is sometimes used, which is the profile that gives a uniform pressure on a signal-free disc at 65 dyne tracking force. This profile may be increased by a "curvature factor" S.

General Features of the Calculated Deformation

The type of signal for which pickup is most affected by the details of elastic deformation is the low-amplitude audio-only signal played by a shoe of low curvature. The shoe is here in contact with the disc surface along almost its whole length, so that there is no modulation at the electrode-stylus contact plane. It might be expected that this would lead to a very low value of capacitive pickup from the electrode. Further examination shows, however, that the variation of pressure as the signal passes under the shoe produces an enhanced "bulge" close to the pressure edge. The situation is shown in Fig. 4, where the surface configuration is given for 12 successive phases of the signal. The vertical scale is exaggerated by 14 times. The broken line, for reference, gives the dimple that would be produced in a smooth disc. In addition, the pressure profile is plotted. It can be seen that the pressure fluctuates greatly at the edge; in fact, there is a mathematical singularity there, and at the limit of infinitesimal element size the pressure will swing between zero and infinity. It is this sharp change of pressure that causes the local bulge, so that 2000 Å from the edge, the surface modulation is quite large, even though this distance is small compared with the signal wavelength. Since the electrode extends out 2000 Å, the enhanced modulation produces capacitance pickup.

This model therefore provides an answer to the question of why significant audio-only pickup can be obtained even though the signal elements are completely compressed by the stylus. For a standard video signal, the modulation depth is larger, so that complete compression does not occur; in this case the effects of finite elastic distortion are less obvious. In addition, a real stylus-shoe is more rounded near the electrode than is the constant-pressure profile used for Fig. 4. Fig. 5 shows one phase of two similar video + audio

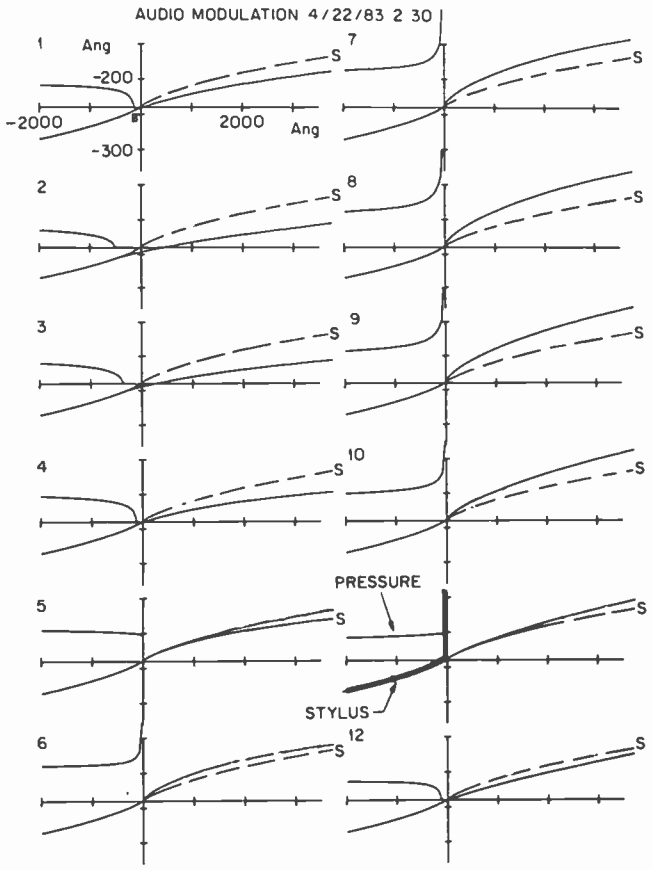


Fig. 4—Surface contour for twelve equally spaced phases of a audio-only signal near the pressure edge. The pressure profile and the stylus outline are identified in phase 11 (vertical/horizontal scale ratio = 14).

signals, with a modulation depth of 850 Å and a wavelength of 1 μm. The "flat-top" signal will be described later. For the phase shown, the disc surface contacts the shoe near the electrode, and the deformation in that region will depend quite strongly on various parameters such as the tracking force. For much of the remainder of the cycle, the electrode-disc spacing will be less sensitive to these parameters. The calculated pickup amplitude will therefore usually not vary greatly from that obtained by more simplified assumptions on the conformation of the surface. What will differ are the second-order effects—harmonic distortion, signal intermodulation, sound-

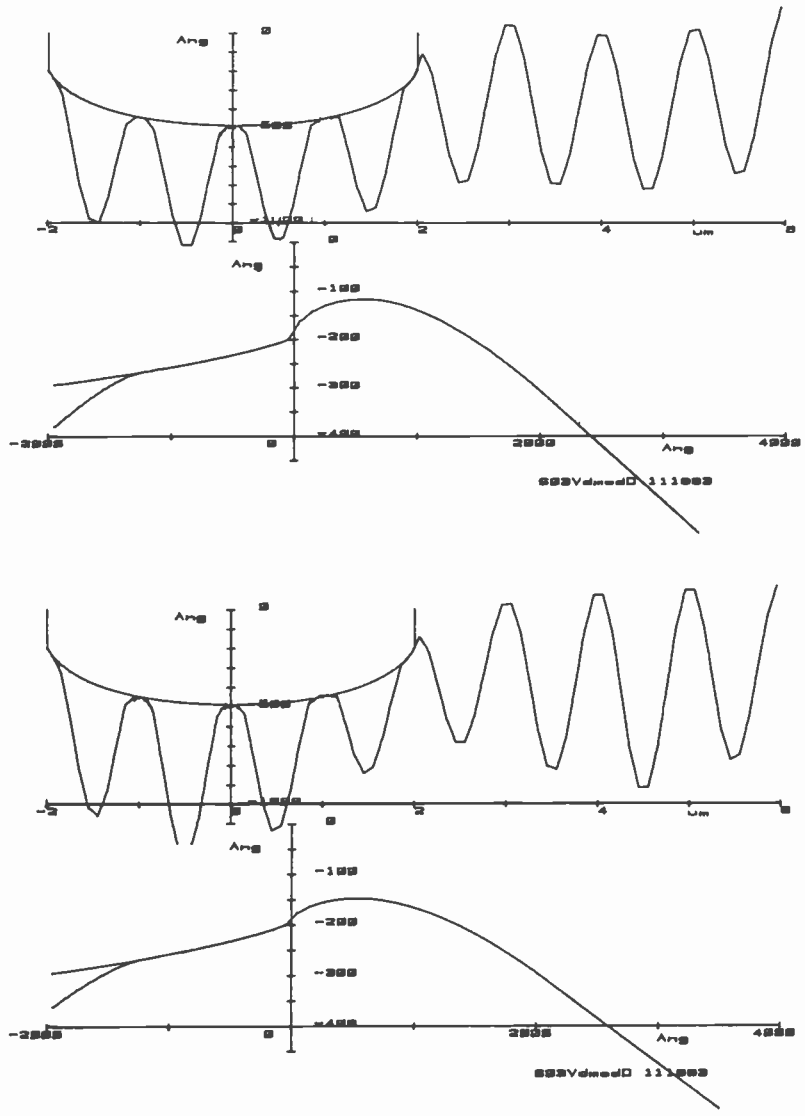


Fig. 5—Surface contours, over the whole shoe and on a magnified scale near the pressure edge for video plus audio signal. Top shows normal signal and bottom "flat top" signal.

beat and the dependence of these on wavelength, tracking force, shoe length, and shoe curvature.

Capacitance Pickup

Having obtained the surface contour, the capacitance between the stylus electrode and the disc (assuming a two-dimensional situation) may be computed exactly using a field distribution program based on Poissons equation. Calculations of this kind have been carried out by N. Binenbaum and R. W. Nosker of these laboratories for the case of a trapezoidal signal form. Significant contributions to the capacitance modulation occur only within about 1 μm of the electrode, so calculations can be confined to this region. In the present paper the electrode is represented as having an angled end and a step-back from the stylus shoe edge, which is a reasonable representation of the form often seen in SEM displays. In addition, the conducting surface of the disc is taken to be displaced below the contact surface due to an intervening dielectric layer (the surface vinyl above the contacting graphite grains) and possibly to an adherent monolayer of oil. Fig. 6(a) shows the dimensions involved.

The exact capacitance calculation involves significant computer time. Again, for the numerous configurations, with the different signals, phases, and electrode shapes needed in this investigation, simplified calculation is desirable. An approximation is used here in which a reasonable estimate is made of the shape of the field lines and of the variation in length of these lines computed as the signal contour passes under the electrode. The contribution to the total capacitance from the element containing the line will then vary inversely as the length of the line, suitably weighted by the dielectric constant.

The dielectric gap is divided into elements bounded by the field lines as shown in Fig. 6(b). The capacitance of each element can then be calculated approximately from its average cross-section area and the length of the lines. Errors in this method will be least in the region where the parallel-plate situation is approximated as in region III of Fig. 6(b), which in fact contributes the majority of the capacitance variation. The shape of the field lines in region III is taken as straight and parallel to the vertical electrode face. In regions I and II, the lines are taken as arcs of a circle for the more distant elements. For elements that are closer than three times the shoe-disc gap on the after side of the electrode, the elements grade from circular arcs to straight lines. An interpolation for both ele-

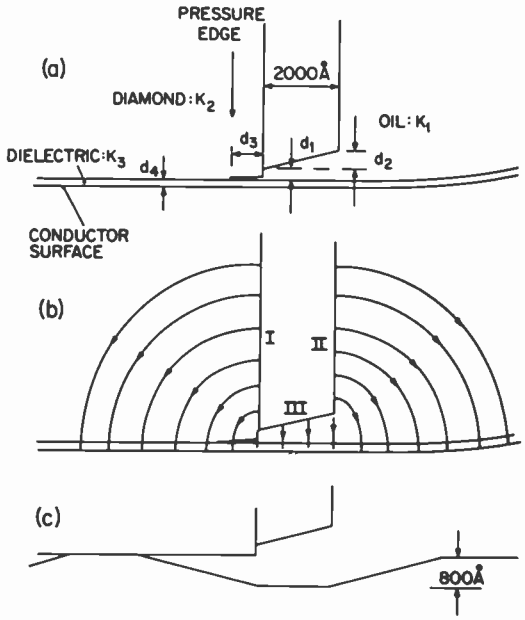


Fig. 6—(a) Identification of electrode and disc spacing dimensions; (b) approximation used for field configuration; and (c) trapezoidal waveform used to check approximate calculations.

ment length and width is made between these two forms for the intermediate elements.

An approximation of this kind can be expected to be only moderately accurate for calculations of absolute capacitance. However, for perturbations of this capacitance, such as are given by the signal form on the disc, a higher accuracy can be expected. An independent derivation of the expected error would be difficult to make, but comparisons with exact calculations for specific forms give a good idea of the general accuracy. Table 1 shows the results of Binenbaum and Nosker for the trapezoidal waveform shown in Fig. 6(c). Using 11 phases of this wave, limited to a 1½ wavelength segment, they derived values of capacitance modulation and its second and third harmonic content for a number of electrode shapes. Table 1 also shows the result of the approximate calculations just described. Fourier analysis of the capacitance values gives fundamental and harmonic pick-up components, which in this paper will be quoted in dB relative to $1aF$ ($10^{-18}F$) RMS (designated dBR) and dB relative to the fundamental, respectively. The approximate results

Table 1—Comparison of Approximate (B) with Exact (A) Electrode-Disc Capacitance Calculations

Electrode and Dielectric Dimensions (Å)				Fundamental (dBR)		Harmonics dB Relative to Fundamental 2nd		3rd	
d_1	d_2	d_4	k_1	A	B	A	B	A	B
200	200	200	2.4	32.9	33.5	−16.6	−16.7	−37.5	−32.0
0	0	200	1	42.1	42.3	−8.8	−8.8	−21.9	−23.3
200	2000	200	2.4	28.2	29.2	−14.7	−14.6		−26.7
200	200	700	2.4	22.7	23.6	−22.7	−21.6		−42.9

agree quite well with the exact values, particularly in the second harmonic content of the pickup. Greater error is seen in the third harmonic at low levels; it is of the order of 1% of the fundamental and is comparable with errors expected from the finite number of sampling points. Agreement is good even when the electrode-disc spacing is as large as 700 Å, in which case the field configuration used in the approximation will differ significantly from the true configuration. The approximate calculations will be used in the remainder of the work to be described.

Calculations of deformed contour and, from them, of capacitance, are made for single sine waves of various amplitudes and wavelengths at each of 12 phases of the wave.

More complex signals, composed of video and audio superposed, are arranged to have an integral number of video wavelengths in an audio signal to minimize calculation of one complete cycle of the total signals. Again, 12 phases of each video cycle are taken. These signals are analyzed by bandpass filtering and computation of the zero-crossing points of the filtered video waveform. Subsequent Fourier analysis then gives the fundamental (i.e., audio-frequency) component of these zero-crossing deviations. Soundbeat, which is the phase modulation of the video carrier pickup by the superposed audio carrier, may then be calculated from the zero-crossings. For a waveform phase-modulated by

$$\cos(\omega_v t + \Delta\phi \cos \omega_a t),$$

soundbeat (in dB) is defined as

$$20\log(5\ \omega_a\ \Delta\phi/\omega_v).$$

The filtering action removes some components that contribute to soundbeat; these are restored by adding a small correction to the value directly calculated from the zero crossings.

Pickup of Single-Frequency Signals

In general, for a constant modulation depth, calculated pick-up tends to increase steadily with the modulation depth and to fall off at long and short wavelengths. The dependence on tracking force is more complex; it may increase, decrease, or go through a maximum. The results to be described have been calculated using a disc Young's modulus of 5.4×10^{10} dyne cm^{-2} and a Poisson ratio of 0.35, near the values measured at 15 MHz by Rehwald and Vonlanthen of RCA Laboratories, Zurich. Fig. 7(a) and 7(b) show the calculated pickup from a sinusoidal signal as a function of tracking force, with separate plots for two different shoelengths (1 dyne is approximately 1 mg force). The electrode is of form E2 [Table 2 and Fig. 6(a)] and the stylus curvature factor is 2. Decrease of pickup with increased tracking force is strongest for long wavelength and

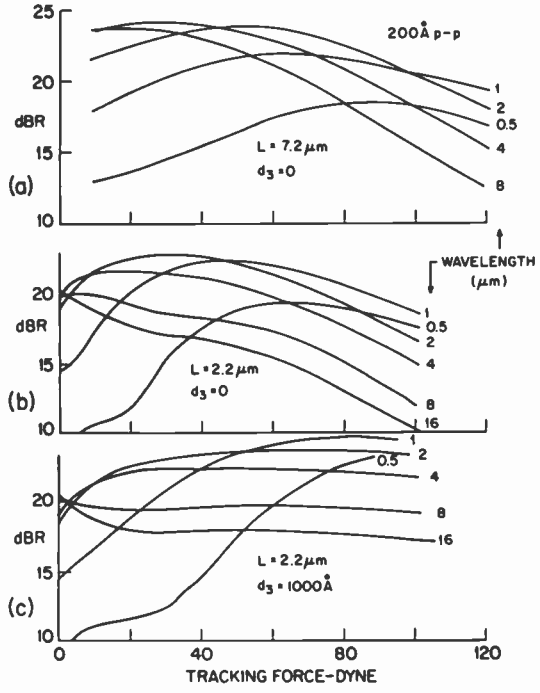


Fig. 7—Effect of tracking force variation on calculated pickup, of a 200 Å p-p signal with wavelengths of from 0.5 to 16 μm. Two shoe lengths L and pressure edge set-back d_3 are shown.

Table 2—Electrode and Disc Dielectric Dimensions (See Fig. 6)

Electrode designation	d_1	d_2	d_3	d_4
E1	200	500	0	200
E2	300	1000	0	200

small signal depth, that is, when all angles to the horizontal are minimized. The increased tracking force then compresses the signal contour. The opposite effect of increased signal with increasing tracking force is seen for short wavelength and large signal depth, where the dominant effect is to bring the electrode closer to the disc surface rather than to compress the contour. Changes of electrode form, or thickness of disc surface dielectric layer, will change the magnitude of the pickup amplitude but, within limits, do not have a large effect on the shape of the curves. A set-back of the pressure edge or an increase of shoe length both tend to decrease the effect of tracking force variation. Fig. 7(c) shows the marked flattening that occurs with a pressure-edge set-back of 1000 Å. A set-back of this magnitude would be difficult to detect experimentally; the difference between Fig. 7(b) and 7(c) therefore illustrates the caution that should be exercised in comparing calculation with experiment.

Curves similar to those of Fig. 7 can now be used to obtain the pickup at constant tracking force. A typical variation with wavelength is shown in Fig. 8 for a signal depth of 200 Å p-p. Calculated curves are shown as dashed lines and are compared with measured curves (solid lines). The electrode gap value has been adjusted to give approximately the same pickup magnitude as measured; these values are in the 300 to 800 Å range and correspond roughly to those estimated from SEM observations in each case. A "standard" shoe shape with a curvature factor of 2 has been used. The tracking force is 60 dyne for the top two curves and 40 dyne for the lower curve. These adjustments from the total applied tracking force of 65 dyne have been made to account for a proportion of the force supported by oil and debris outside the shoe area, guessed to be 10%. For curve C in Fig. 7, the shoe width is 40% greater than that for curves A and B. In addition the shoes have been given a slight tilt to stimulate a possible wedge action of the oil film, as discussed before; the prow end is 150 Å lower than the electrode end of the shoe.

The curves show a fall-off at both short and long wavelengths. At the longer wavelengths the signal profile can conform more closely

to the shoe and so a fall-off can be expected, beginning at wavelengths comparable to the shoe length. At short wavelengths, the effective electrical aperture of the electrode capacitance becomes comparable with the wavelength and this reduces pickup. However, for the shorter shoes, the calculated fall-off is less than measured. One factor that could change the fall-off rate is the assigned electrode gap. This does not seem to be responsible, however, since a change of gap that produces a large deviation of absolute value of pickup from that measured does not produce a significantly better shape match. A second possible factor is the skewing of the electrode edge line from the signal peak ridge. This also cannot explain the difference, since a skew sufficient to produce a 3dB change in output would also produce a sharp change in the slope of the pickup versus wavelength curve, which is not seen in the measurements. Different shoe shapes and tracking force, within limits that reasonably correspond to measurement, do not produce a significantly better fit to measurement than seen in Fig. 8.

Measurements of dependence of pickup on signal amplitude, using a shoe of 5.2 μm overall length, are shown in Fig. 9. Comparable calculations for wavelengths of 1 and 8 μm are shown as dashed lines, using a shoe of profile matched to SEM measurements and with a corresponding effective length of 4.6 μm. The stylus is given an 80 Å tilt and 60 dyne tracking force. The most noticeable characteristic of the measurements is the flattening of the curve at higher amplitudes, which occurs more strongly at the shorter wavelengths. A similar effect is seen in the calculated curves at 1 μm wavelength. The linearity of the 8 μm calculated curves, however, is at variance with the slight curvature measured at both 4.2 and 13.3 μm, at which wavelengths the original signal is believed to be well-calibrated. Various stylus and electrode shapes have been tried, but all give curves close to linear; this discrepancy also has not been resolved within the framework of the model.

Comparison of the measured harmonic content of pickup with calculations using the same shoe profile as in Fig. 9 is given in Fig. 10. The electrode gap again was adjusted to give the same magnitude of the pickup amplitude as measured. The calculated magnitude of both second and third harmonic is of the same order as measurement at higher signal amplitude; at low amplitudes, the calculated harmonics do not fall off as rapidly as measured. Below −30 dB, calculation is not expected to give accurate results. In some regions, a rapid change of harmonic content with tracking force was calculated, in which case harmonic levels were averaged over the 50- to 70-dyne range.

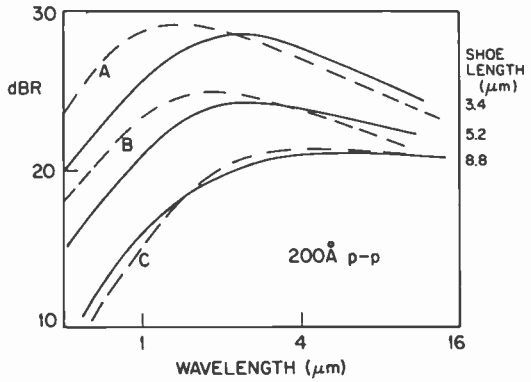


Fig. 8—Calculated variation (dashed lines) of pickup of a 200 Å p-p signal with wavelength compared with measurement (solid curves).

Results of several sets of measurements of pickup amplitude and harmonic content as a function of tracking force are shown in Fig. 11 as solid lines and dot-dash lines. Both audio (97 Å p-p, approx. 1 μm wavelength) and video (800 Å p-p, approx. 7 μm wavelength) signals were measured. The scale for the dot-dash lines is arbitrary since no absolute capacitance calibration was made for this set of

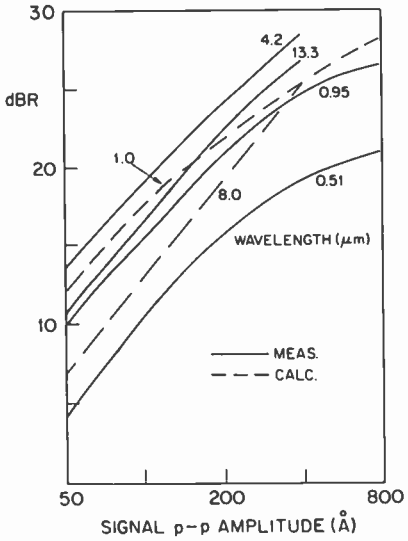


Fig. 9—Amplitude dependence of pickup at different wavelengths in μm (solid curves are measured values, dashed lines are calculated values).

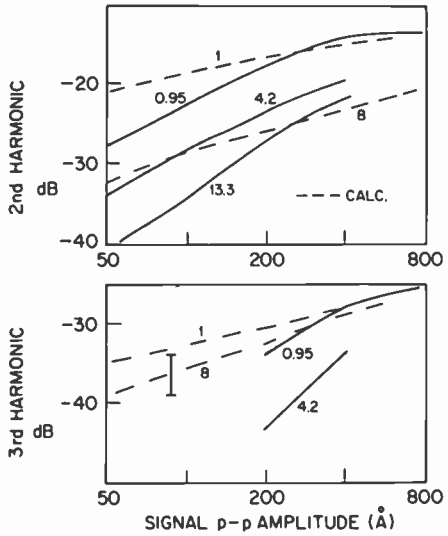


Fig. 10—Harmonic content of pickup as a function of signal amplitude (solid curves are measured values, dashed lines calculated values). Wavelength in μm on curves.

measurements; the video response has been placed close to that of the other data. The tracking-force dependence of the audio response shown by one of the solid lines can be seen to differ significantly from that of the other curves, more in fact than the error bars which show the data point scatter. The open circles shown are single points from another set of measurements at 65 dyne. It is evident that there is a considerable spread in the experimental data, produced by factors whose nature has not been determined.

Corresponding calculations of pickup amplitude and second harmonic dependence on tracking force are shown in Fig. 11 as dashed lines. An effective shoe length of 4 μm has been used, together with an electrode gap of about 500 Å, which gives an audio pickup amplitude within the range measured. The calculated tracking-force dependences, the audio/video ratio, the second/third harmonic ratios, are all consistent with the type of measured behavior. Variations in tracking-force dependence of measured audio pickup are consistent with the variations calculated in Fig. 7 for different shoe set-back values. The comparison therefore gives some plausibility to the model, even if a close correspondence cannot be made.

Fig. 12 shows the tracking-force dependence of the 5 MHz, 2 MHz, and 8 MHz signals. These measurements are shown as dashed lines

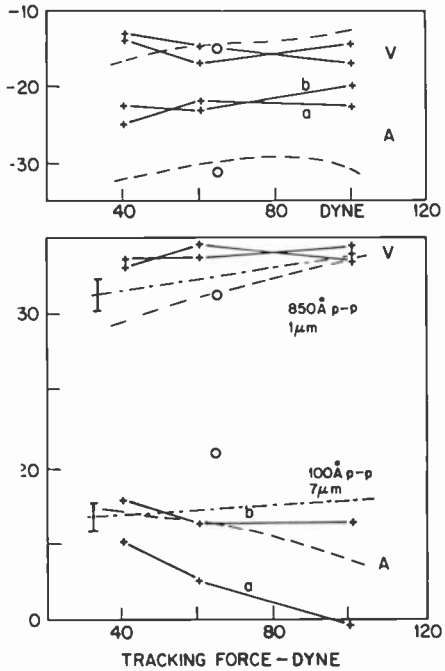


Fig. 11—Pickup amplitude and harmonic content as a function of tracking force for video (850 Å, 1 μm wavelength) and audio (100 Å p-p, 7 μm wavelength) signals. The solid lines, dot-dash lines, and circles are measured values; the dashed lines are calculated values. An arbitrary dB zero is used for the dot-dash curves.

with error bars to indicate scatter of the measured points; the absolute amplitude scale for the whole set again, is arbitrary. The solid curves show the result of present calculations with the same shoe as in Figure 11; the variation of the calculated curve slope with signal frequency is consistent with that measured.

Soundbeat

Details of the mechanism generating soundbeat are quite subtle and the way in which it may vary with various pickup parameters is not intuitively obvious. Understanding of the soundbeat mechanism is gained if two separate components of the change of capacitance, before and after the stylus-electrode plane, are considered separately. It then becomes apparent that the soundbeat of each individual component is less than that of the combination. There is,

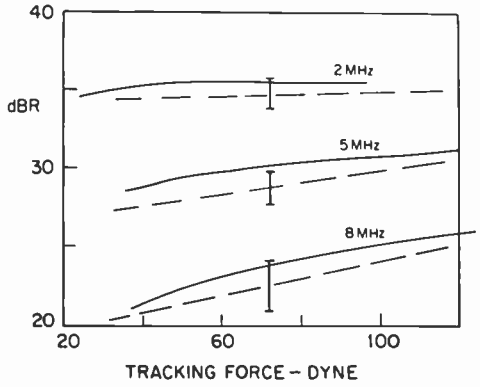


Fig. 12—Dependence of pickup on tracking force at three signal frequencies (5 MHz = 1 μm wavelength). Solid curves are for calculated values and dashed lines for measured values (arbitrary dB zero). Measurement error is indicated.

however, an amplitude modulation at the audio frequency that is different for the two parts, being less for the forward part which is compressed under the shoe. Since there is a substantial phase-shift between the two parts, the amplitude modulation translates to a phase modulation in the total pickup. This is to be contrasted with the situation of a stylus on a rigid, undeformed signal contour. Here the amplitude modulation of the two parts is of similar magnitude; when they are summed, the net phase modulation is small. Fig. 13 shows the after and before parts (A and F) and the total (T) calculated capacitance modulation for a deformed contour; the larger amplitude modulation of the after part can be seen. The upper trace (S) is the bandpass filtered signal, where the audio component and the harmonics have been removed from the T trace. The soundbeat is too small to be visible on this scale.

Table 3 gives the modulation for a typical signal-stylus combination with 85 Å p-p audio, for normal disc elasticity and tracking force and also for a rigid contour (equivalently obtained by setting a zero tracking force). The electrode gap has been adjusted to give approximately similar pickup amplitude in each case. With the normal situation, the video amplitude modulation depths are 0.2 dB and 0.9 dB before and after the electrode, respectively. This small difference, together with the video phase shift of 0.187 of a cycle, is the primary cause of the soundbeat, which is 4.6 dB greater than the average of the two parts. In comparison with this, for the rigid contour, the amplitude modulation of both parts of the video signal is the same (1.5 dB), so there will be no extra component

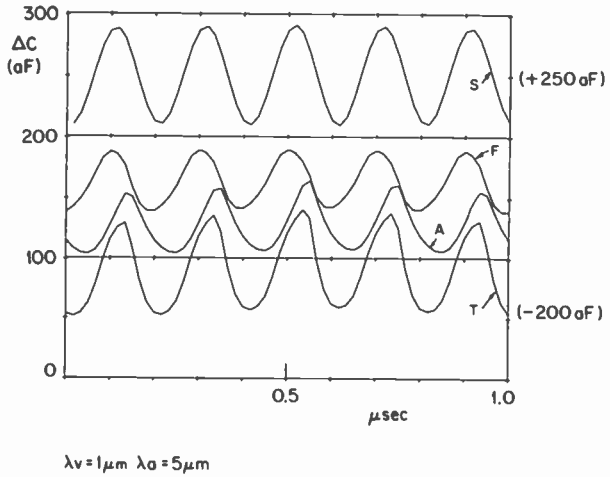


Fig. 13—Calculated components of total pickup from before (F) and after (A) the stylus/electrode plane. The total signal (T = A + F) is shown bandpass filtered in the upper trace (S).

added to the soundbeat of the individual parts. The phases of the individual parts, however, are opposed, being 0.40 and −0.11 of a cycle respectively. As a consequence, the total computed soundbeat is reduced to −48.6 dB. This difference value will be sensitive to electrode configuration.

When the contour is rigid, the phase modulation of the after part

Table 3—Contribution of Forward and After Parts to the Total Soundbeat (SB) and Pickup Amplitude for Deformed and for Rigid Contours

Part of Stylus Plane	Contour	
	Deformed	Rigid
After		
SB (−dB)	31.9	35.5
(Phase)	0.46	−0.11
Video (dBR)	25.8	27.5
(dB Modulation)	0.9	1.5
(Phase)	0.660	0.649
Forward		
SB (−dB)	31.3	31.7
(Phase)	0.46	0.40
Video (dBR)	25.3	25.3
(dB Modulation)	0.2	1.5
(Phase)	0.473	0.446
Total		
SB (−dB)	27.0	48.6
Video (dBR)	30.0	30.6

is opposite to that of the forward part, because a motion of the disc surface toward the stylus causes greater relative contribution from the capacitance elements near the center plane. The phases of video pickup from the two parts therefore move toward each other. It may then be questioned why this does not happen when the contour is normally deformed. Table 3 shows that the video phase of the after part moves back rather than forward as the surface approaches the electrode at the peak of the audio. The reason is not clear. It must be supposed that the contour distorts in such a way that a compensating phase shift away from the electrode is produced.

Calculations of soundbeat are shown in Fig. 14 for a shoe of effective length 3.5 μm and a profile matched to SEM measurement of a normal shoe. These are compared with scattered values measured using a number of standard styli with similar shoe lengths, averaging about 4.5 μm overall, and a test disc with an audio amplitude of 85 Å p-p and a video wavelength of approximately 1 μm. The three calculated points in Fig. 14 correspond to three different electrode gaps. The points are seen to lie well within the scatter of measurements but about 1 dB lower than the mean. The drop of soundbeat with increased pickup is consistent with measurement, although the scatter is too great to give this any real significance.

Two sets of measurements of the dependence of soundbeat on signal wavelength are shown in Fig. 15(a). Both sets used test discs in which the signal frequency remains constant, so that the wavelength varies as the playing radius changes. In all cases the audio amplitude was 100 Å p-p and the video 850 Å p-p; the video fre-

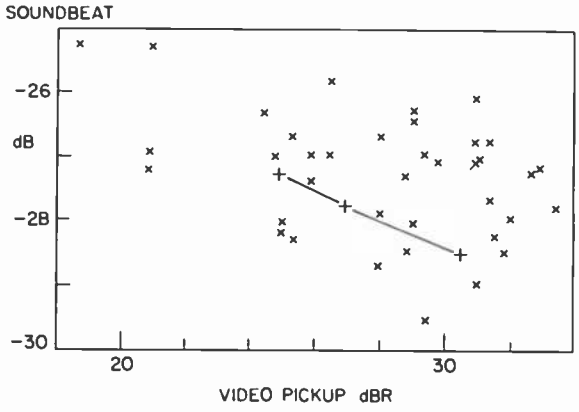


Fig. 14—Calculated soudbeat (solid line) as a function of pickup change caused by electrode-gap variation compared with measurements using a number of standard styli.

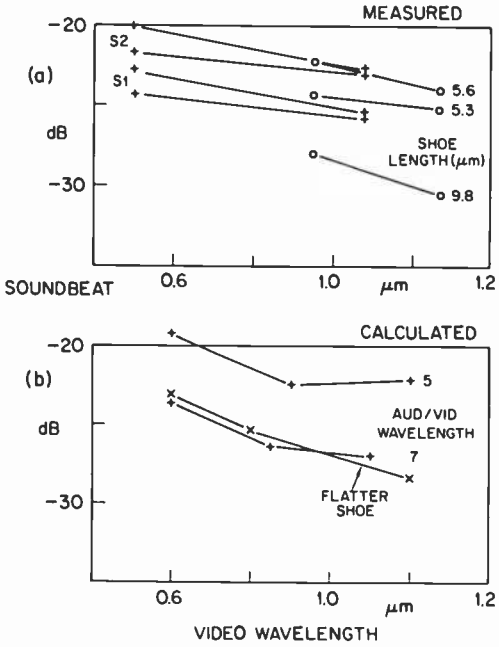


Fig. 15—Soundbeat for normal signal as a function of wavelength: (a) measured values of soundbeat for different shoe lengths (values shown by crosses are from R. W. Nosker; circles are from this study (uncertain dB zero)); (b) calculated values for video output level similar to measured values for different audio/video wavelength ratios.

quency was 5 MHz and the audio less by a factor of 7 (S1 signal) or of 5.6 (S2 signal). The steeper lines of the measurements shown by crosses were made with a flatter stylus shoe. Lines shown by circles (S1 signal only) include a small estimated correction for the zero offset of the whole group.

Calculated curves corresponding to these measurements are shown in Figure 15(b). Again, the stylus shoe profile was taken as approximating those seen in SEM and the electrode spacing adjusted to give the magnitude of video pickup that was measured. The shorter-wavelength audio signal in the calculations had only 5 times the video wavelength and so does not correspond exactly to the S2 factor of 5.6. Measurement and calculation are in reasonable agreement, although there is still a tendency for calculated soundbeat values for the normal length shoe to be about 1 dB lower than measured; changes of stylus profile, shoe length, or electrode

spacing within a reasonable range did not increase soundbeat significantly above values in Fig. 15(b).

The points marked "flatter shoe" were calculated for a shoe with curvature 0.7 of that used in the other curves; it shows a stronger wavelength dependence, as in the measurements. Calculations for an 8 μm shoe show little change of soundbeat from the normal-length shoe if the shoe is level, which is in contrast to the 7 dB drop measured for the 9.8 μm real shoe compared to a 5 μm real shoe (plots c versus a and b in Fig. 10(a)). The reason for this may be the higher sensitivity of the longer shoe to shoe tilt, that is, to a slightly prow-down condition. A further calculation using a shoe that is 200 Å prow-down, corresponding to the tilt produced by a rise in record level of only 4 mils, shows a dramatic drop in soundbeat, on the order of 10 dB. This is primarily due to the drop of stylus pressure near the electrode as the shoe tilts to a prow-down condition. For a normal length shoe, the effect is calculated to occur only for much larger tilts.

Finally, the effect of a flat-top signal may be noted. Soundbeat can be reduced by modifying the waveform so that all signal contour peaks are at the same level:

$$A = A_o \cos k_a x + (A_v - A_a \cos k_a x) \cos k_v x$$

Table 4 shows the soundbeat calculated for this waveform, for an S1 audio signal of 85 Å p-p amplitude ($2A_o$), and a shoe of curvature factor 1.5 compared with that for a normal signal. The reduction of soundbeat is substantial: 7.5 dB for electrode E2 (Table 2), which gives a normal video signal output level, and 9.3 dB for E1, which is close-spaced and gives a signal about 4 dB above normal. Measurements with the same three styli as Fig. 15(a) for a flat-top signal with a nominal 100 Å p-p audio component gave soundbeat values that were lower and had a dependence on video wavelength in the opposite direction to that of a normal signal. Calculated

Table 4—Effect of Flat-top Signal on Soundbeat (SB) for Electrodes E1 and E2 (See Table 2)

	E1		E2	
Signal	SB (−dB)	Video (dBR)	SB (−dB)	Video (dBR)
Normal	27.7	33.5	28.0	29.4
Flat	37.0	33.5	35.5	29.4

values have a similar dependence on wavelength but a magnitude that represents an approximately 3 dB greater reduction from the normal signal. This difference may be due to the way in which the disc signal is produced. Even though a flat-top electrical signal was fed to the cutter head that made the original master, the cutting process may be sufficiently nonlinear to reintroduce some added audio signal.

Conclusions

The general agreement of the calculations with measurement confirms the broad validity of the contour deformation model of stylus pickup. Many of the apparent discrepancies can be seen to lie within the scatter of experiment or to involve unknown factors, such as the linearity of the signal-cutting mechanism. Agreement is best when an electrical calibration is not involved, as in the dependence on tracking force. A more extensive and correlated set of measurements would be required to elucidate the effect of all stylus and disc parameters and obtain a complete understanding of the limitations of the model.

Acknowledgements

The author is indebted to C. H. Anderson, R. W. Nosker and J. R. Matey for a number of informative discussions on the basic physical concepts of this analysis; to R. W. Klopfenstein for showing the method of solution for the pressure distribution; and to J. J. Gibson for discussion of alternate analyses and for the design of the flat-top signal geometry. The various measurements quoted have been made, at varying locations within RCA, by D. Brigham, J. E. Economou, J. R. Matey, R. W. Nosker, J. G. Pecorari, and E. D. Simshauser, with the assistance of G. R. Auth and J. F. McLaughlin; their contribution to this paper is gratefully acknowledged.

References:

[1] J. K. Clemens, "Capacitive Pickup and the Buried Subcarrier Encoding System for the RCA VideoDisc," *RCA Rev.*, 39, p. 33, March 1978.
[2] S. P. Timoshenko and J. N. Goodier, *Theory of Elasticity*, 3rd Edition, McGraw-Hill Book Co., p. 402.

Some Applications of z-Transforms and State Variables to Feedback Shift Registers

Harry Urkowitz
RCA Government Systems Div., Moorestown, NJ 08057

Abstract—After a brief review of the elementary properties of binary linear recursive sequences, two types of feedback shift register for their generation are introduced. One is the simple feedback shift register generator (SSRG) and the other is the multiple return, or modular, shift register generator (MSRG). With appropriate feedback connections, both types can generate the same sequence.

The values in the storage elements of a feedback shift register constitute the state vector of the register. The succession of states describes completely the autonomous (self) behavior of the register. A particularly apt method of analyzing the succession of states involves discrete state variables. Through their use and the use of *z*-transforms, equivalences are established between an SSRG and an MSRG to yield the same sequence, between parallel combinations and a single large register, between cascade combinations and a single larger register, and some equivalences are established involving complemented sequences including alternating complements and generalized complements.

1. Introduction

In this paper we consider two forms of feedback shift registers that generate binary linear recursive sequences. In particular, we consider conversion of states from one form to the other and cascade and parallel combinations. These equivalences are obtained by the application of *z*-transforms and by the state variable approach.

Both *z*-transforms and state variables have been applied to feedback shift registers before, particularly by Huffman[1,2] and by Pe-

terson,[3,4] but there are some differences in our approach. Huffman introduced the delay operator D which is formally equivalent to $1/z$ of the bilateral z-transform. In this paper we use the unilateral z-transform because initial conditions are more readily inserted. The autonomous behavior is readily studied by means of state variables (particularly the use of the state transition matrix) used by (among others) Elspas[5] and Birdsall and Ristenbatt,[6] although the usual modern terminology was not used by them.

Of particular interest to us are the two forms of feedback shift register, called by Birdsall and Ristenbatt the "simple" shift register generator (SSRG) and the "modular" or "multiple return" shift register generator (MSRG). The SSRG uses several cascaded delay or storage elements and has a single feedback reentry point. The MSRG has multiple reentry points in the feedback paths so that binary adders are required between some of the storage elements. While appropriate connections of the two forms will produce identical binary sequences, the states (i.e. the contents) of the two forms, as the elements of the output sequence are generated, will be different, in general. There are two reasons for treating both forms:

(1) The successive states of an n-stage SSRG are successive N-digit portions of the output sequence, i.e., the contents of the SSRG appear in the output sequence.
(2) The successive states of an n-stage MSRG may be identified with the elements of a Galois (i.e., finite) field of characteristic 2; i.e., we may say that an MSRG performs "field counting". This is pointed out by Peterson[3,4] and by Berlekamp.[7]

Peterson points out that there is a one-to-one correspondence between the states of the two types of shift register, but formulas relating the two are not given. By combining the use of state variables and unilateral z-transforms, explicit formulas are derived here for conversion from an SSRG state to an MSRG state.

In this paper, we first apply the one-sided z-transform to a linear recursive sequence and show how initial conditions enter. These initial conditions are identified with the state of an SSRG that generates the linear recursive sequence.

We then couple z-transforms with state variables. Transfer functions are derived as rational fractions in z. Together with the state transition matrix describing the autonomous behavior, z-transforms are used to derive the equivalence between MSRG and SSRG states.

Next, using the coupling mentioned above, we obtain the behavior of a series and a parallel combination of either type of shift register,

together with derivations of equivalent initial states. Among the results obtained is the extraction of the component sequences from the sequence produced by a cascade.

Last the z-transform is used to show some interesting algebraic properties of complements, alternating complements, and generalized complements of binary sequences.

It is believed that almost all of the results beginning with Sec. 5 are new except, possibly, for the application to dual-state modulation of the z-transforms of singly complemented sequences. The material on generalized complements has not yet, as far as the author knows, found application to modulation schemes.

For the reader's convenience, the definition of symbols used in this paper are given in Appendix B.

2. Linear Recursions and Their z-Transforms

A *recurrence* or recursion is a relationship that describes how each element of a sequence depends upon the previous elements. A *linear recurrence* refers to a linear combination of past elements. Thus, let $\{y(k)\}$, $k = 0, 1, 2, \ldots$, represent a sequence of elements $y(k)$. Then a *linear recurrent* sequence, or *linear recursive* sequence, satisfies an equation of the form:

$$\begin{aligned} y(k+n) &= f_0 y(k) + f_1 y(k+1) + \ldots + f_{n-1} y(n+k-1) \\ &= \sum_{j=0}^{n-1} f_j y(k+j) \\ &= \sum_{j=k}^{n+k-1} f_{j-k} y(j), \; k = 0, 1, 2, \ldots \end{aligned} \qquad [1]$$

where the coefficients f_j are elements of a field and do not depend on k. For the case of integers modulo 2, the f_j will be 0 or 1 and the addition will be modulo 2. Eq. [1] refers to an nth order sequence.

It may be seen that the equation that is satisfied by a linear recursive sequence is a linear difference equation and can be analyzed with all of the tools available for such equations, including z-transforms.

Suppose we rewrite Eq. [1] in the following way:

$$\sum_{j=0}^{n} f_j y(k+j) = 0, \; k \geq 0, \qquad [2]$$

where it is understood that $f_n = 1$ and all arithmetic is modulo 2.

Eq. [2] is a homogeneous linear difference equation with constant coefficients. Being homogeneous, it has a solution or solutions that depend only on initial conditions; no driving force is implied, except for the initial conditions. Thus, when we consider physical devices for the generation of sequences that are solutions of Eq. [2], we will consider their undriven or "autonomous" behavior. The sequence that comes out of such a network naturally consists of a set of digits arriving sequentially in time. Thus, the network may be known as a "sequential" network. Therefore, we may say that we are studying linear, autonomous, sequential networks.

The shift register generators we study have periodic sequences and the maximum period obtainable from an n-stage feedback shift register is $2^n - 1$. The first n digits of the sequence are the initial conditions and must be assigned. However, in the sequence, because of the recursive relationship of Eq. [2], any n digits may be considered as the initial conditions for the rest of the sequence.

More insight may be obtained from the z-transform of the sequence.[8] The (one-sided) z-transform of a sequence $\{y(k)\}$, $k \geq 0$, is defined by

$$Z\{y(k)\} \triangleq Y(z) \triangleq \sum_{k=0}^{\infty} y(k)\, z^{-k}. \qquad [3]$$

Eq. [3] can be applied to Eq. [1] by rearranging it. Assuming that the f_j are integers in the field modulo 2, we have, after some manipulation,

$$Y(z) = \frac{\sum_{j=1}^{n} f_j \sum_{i=0}^{j-1} y(i) z^{j-i}}{f(z)} = \frac{\sum_{j=1}^{n} z^j \sum_{i=j}^{n} f_i y(i-j)}{f(z)}$$

$$= \frac{\sum_{j=1}^{n} z^j \sum_{i=j}^{n} y(n-i) f_{n-i+j}}{f(z)}, \qquad [4]$$

where the polynomial $f(z)$ is

$$f(z) = f_0 + f_1 z + f_2 z^2 + \ldots + f_n z^n. \qquad [5]$$

$f(z)$ is called the "characteristic polynomial" of the sequence.

3. Transfer Function and Impulse Responses

In this section we determine the transfer function and the state transition matrix of both types of feedback shift register. The es-

sential elements of binary sequential networks are the binary adder, the multiplier (having the value 0 or 1), and the delay or storage element.

These basic elements are combined to form feedback shift registers as shown in Figs. 1 and 2. Both the plus sign and the subscripted sigma indicate addition modulo 2. Fig. 1 shows two representations for the SSRG, while Fig. 2 shows two representations for the MSRG. For either generator, the *state* component or value of any particular stage is taken as the value that would be obtained if that stage were sampled. Therefore, the output $y(k)$ is the content of the last stage or storage element at shift k. The feedback weights, in both cases, are either 0 or 1. When necessary, we distinguish the state, or set of contents, of each type of generator by different subscripts: $x_{sj}(k)$ for the SSRG and $x_{Mj}(k)$ for the MSRG. Note that the weights for both types of generator have been labeled to correspond, so that the external behavior is to be the same.

The impulse sequence response or, simply, the impulse response of either form, is obtained by setting the input $u(k)$ to the unit pulse:

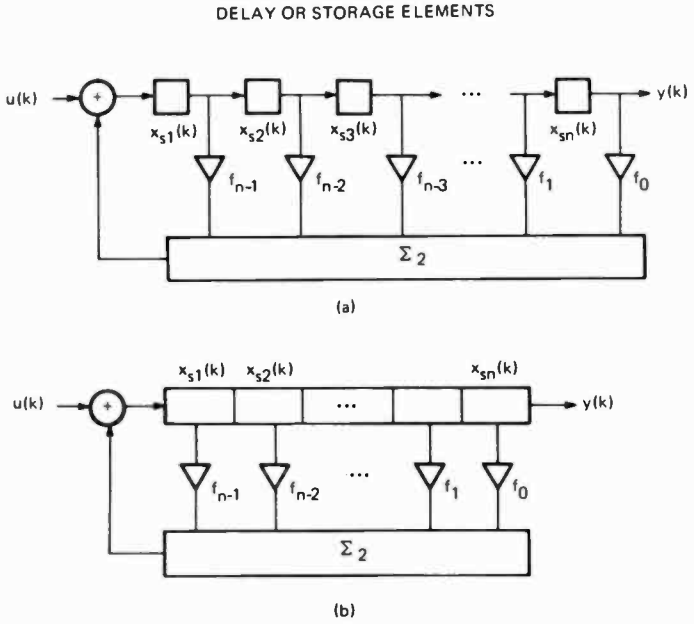


Fig. 1—Two representations for the simple feedback shift register (SSRG). At shift k, the values stored in the stages are $x_{s1}(k)$, $x_{s2}(k)$, . . . , $x_{sn}(k)$. The weights f_i are 0 or 1. At shift k, the output $y(k)$ is taken as $x_{sn}(k)$.

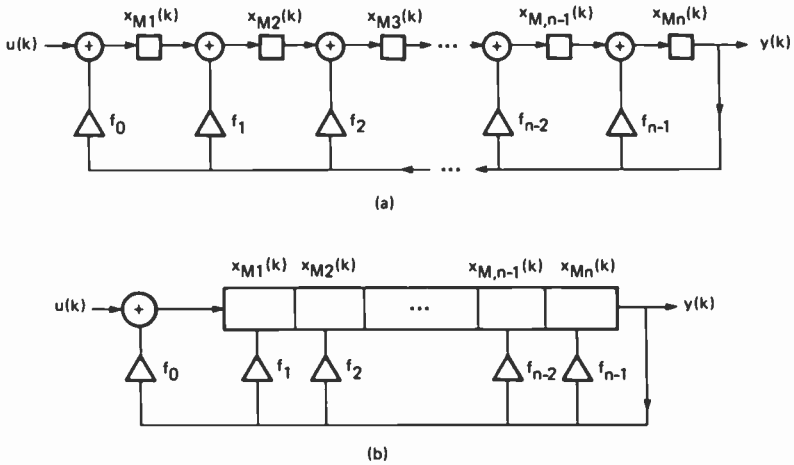


Fig. 2—Two representations for the modular shift register (MSRG). At shift k, the values stored in the stages are $x_{M1}(k)$, $x_{M2}(k)$, . . . , $x_{Mn}(k)$. The weights f_i are 0 or 1. At shift k, the output $y(k)$ is taken as $x_{Mn}(k)$. The weights f_i are labeled to correspond with those of Fig. 1 to provide the same external behavior.

$$u(k) = \delta_0(k), \qquad [6]$$

where $\delta_0(k)$ is 1 for $k = 0$ and 0 for $k \neq 0$. No initial conditions exist within the generator. The transfer function is the z-transform of the impulse response and may be found by direct analysis. The following terminology will be used:

$h_s(k)$ = impulse response of the SSRG
$h_M(k)$ = impulse response of the MSRG
$H_s(z)$ = transfer function of the SSRG
$H_M(z)$ = transfer function of the MSRG

Straightforward analysis, with the aid of Fig. 1, shows that

$$H_s(z) = \left(\sum_{j=0}^{n} f_j z^j = 1/f(z) \right)^{-1}, \qquad [7]$$

with f_n taken as unity. Similarly, with the aid of Fig. 2, it is found that

$$H_M(z) = 1/f(z). \qquad [8]$$

Of course, we anticipated the equality of $H_s(z)$ and $H_M(z)$ by our labeling of Figs. 1 and 2.

It is easy to show that $h_s(k)$ or $h_M(k)$ satisfies the recurrence described by Eq. [2].

4. State Variable—Autonomous Behavior

Now let us turn to the application of state variables as a means for describing the autonomous behavior. An appropriate reference for the application of state variables is Cadzow.[8]

A system may be characterized by an internal *state vector,* whose components are *state variables.* The state variables may be chosen in many ways, but we shall choose the values of the numbers stored in the shift register as state variables. As time passes, the *state* of the shift register will change and the way in which the state vector changes with time and the way in which the outputs vary make up the state space description or state variable description.

Our interest at the moment is in the autonomous behavior; there will be no input so the sequence $u(k)$ is zero. We concentrate on the state vector $\mathbf{x}(k)$:

$$\mathbf{x}(k) = [x_1(k) \ldots x_n(k)]^T \qquad [9]$$

and the scalar *output sequence* $y(k)$. The index k is a time index and is an integer. No subscript indicating the SSRG or the MSRG is used at this point because the discussion is general so far. When it becomes necessary to distinguish between the two forms of generator, the subscripts s and M will be used, as in Figs. 1 and 2.

For our purpose, the autonomous system dynamics are described by a set of first-order difference equations, called the state vector equation and the output equation:

$$\mathbf{x}(k+1) = \mathbf{A}\,\mathbf{x}(k) \qquad [10a]$$

$$y(k) = \mathbf{C}\,\mathbf{x}(k) \qquad [10b]$$

$\mathbf{A}$ is an $n \times n$ matrix called the *system matrix* and $\mathbf{C}$ is a row vector of n components. In fact, examination of Figs. 1 and 2 shows that $y(k) = x_n(k)$, so that

$$\mathbf{C} = [00 \ldots 1]. \qquad [11]$$

Another quantity of importance is the *state transition* matrix $\Phi(k)$, defined as follows:

$$\Phi(k) = \mathbf{A}^k, \ \Phi(0) = \mathbf{I}. \qquad [12]$$

Using $\Phi(k)$, the solution of Eq. [12] is the pair

$$\mathbf{x}(k) = \Phi(k)\,\mathbf{x}(0) = \mathbf{A}^k\,\mathbf{x}(0) \qquad [13]$$

$$y(k) = \mathbf{C}\,\Phi(k)\,\mathbf{x}(0) = \mathbf{C}\,\mathbf{A}^k\,\mathbf{x}(0) \qquad [14]$$

The system matrices for the SSRG (subscript s) and for the MSRG (subscript M) are:

$$\mathbf{A}_s = \begin{bmatrix} f_{n-1} & f_{n-2} & f_{n-3} & \cdots & f_1 & 1 \\ 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ & & \vdots & & & \\ 0 & 0 & 0 & \cdots & 1 & 0 \end{bmatrix} \qquad [15]$$

$$\mathbf{A}_M = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & 1 \\ 1 & 0 & 0 & \cdots & 0 & f_1 \\ 0 & 1 & 0 & \cdots & 0 & f_2 \\ 0 & 0 & 1 & \cdots & 0 & f_3 \\ & & & \vdots & & \\ 0 & 0 & 0 & \cdots & 1 & f_{n-1} \end{bmatrix} \qquad [16]$$

where it is presumed that $f_0 = f_n = 1$.

There are two aspects to the autonomous behavior. One is the succession of states; this is given by Eq. [13]. The other aspect is the output sequence for a given initial state vector. This is given by Eq. [14] with the system or A matrix given by Eqs. [15] and [16] for the two types of shift registers. Eqs. [13] and [14] can be viewed as the solution or description of the autonomous behavior. However, it is useful to express Eq. [14] in terms of z-transforms; these will be needed in the sequel when conversion from one generator to the other is considered. Specifically, we seek the z-transform of the output sequence $y(k)$ for the two types of shift register generator.

Using the subscripts s and M, respectively, for the SSRG and MSRG, it is shown in Appendix A that the z-transforms of the output sequences in the two cases are given by

$$Y_s(z) = \frac{\sum_{j=1}^{n} z^j \sum_{i=j}^{n} x_{si}(0)\, f_{n-i+j}}{f(z)}, \qquad [17]$$

$$Y_M(z) = \frac{\sum_{j=1}^{n} x_{Mj}(0) z^j}{f(z)}. \qquad [18]$$

In the next section, we equate Eqs. [17] and [18] in order to show

equivalent or corresponding states of the two forms of shift register generator. However, there is an additional relationship between the components of the state vector of the SSRG which will be found useful in what follows. If we examine Fig. 1, we see that $x_{s1}(k)$ is obtained from the weighted sum of the components from the previous shift $k - 1$. Thus, it follows that

$$x_1(k + 1) = \sum_{j=0}^{n-1} f_j \, x_{n-j}(k) = \sum_{j=1}^{n} x_j(k) \, f_{n-j}. \qquad [19]$$

5. Equivalent or Corresponding States

Let us consider two shift register generators, one an SSRG and the other an MSRG. Let them both have the same characteristic polynomial $f(z)$. Let us demand the same output:

$$Y_s(z) = Y_M(z) = Y(z). \qquad [20]$$

What will be the corresponding initial states to satisfy Eq. [20]? In general, what will be the corresponding states at any value of shift?

To answer these questions, we may equate Eqs. [17] and [18] and arrive at corresponding initial state vectors $\mathbf{x_s}(0)$ and $\mathbf{x_M}(0)$. However, it should be realized that the output following a given state is dependent only on that state and not on the previous state. By considering the shift origin as arbitrary, we get a correspondence or equivalence for any value of shift, k. With these statements in mind, we get:

$$\sum_{j=1}^{n} x_{Mj}(k) \, z^j = \sum_{j=1}^{n} z^j \sum_{i=1}^{n} x_{si}(k) \, f_{n-i+j}. \qquad [21]$$

Equating like powers of z, we get the following equivalence formulas:

$$x_{Mj}(k) = \sum_{i=j}^{n} x_{si}(k) \, f_{n-i+j},$$

$$= \sum_{i=j}^{n} f_i \, x_{s,n-i+j}(k). \qquad [22]$$

It is interesting to note that the constant term f_0 of the polynomial $f(z)$ does not enter into the conversion.

These equations can be expressed in compact form by defining a "state conversion" matrix G as follows:

$$\mathbf{G} = \begin{bmatrix} 1 & f_{n-1} & f_{n-2} & \cdots & f_2 & f_1 \\ 0 & 1 & f_{n-1} & \cdots & f_3 & f_2 \\ 0 & 0 & 1 & \cdots & f_4 & f_3 \\ & & \vdots & & & \\ 0 & 0 & 0 & \cdots & 1 & f_{n-1} \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix} \quad [23]$$

Then Eq. [22] can be expressed as

$$\mathbf{x_M} = \mathbf{G}\,\mathbf{x_s}. \quad [24]$$

The reverse process of expressing the components of $\mathbf{x_s}$ in terms of $\mathbf{x_M}$ can be accomplished by using the inverse of the matrix G, obtained from G by a little tedious algebra, or by a formula which is the counterpart of Eq. [22]. These expressions are

$$\mathbf{x_s} = \mathbf{G}^{-1}\mathbf{x_M}, \quad [25]$$

or

$$x_{sj} = x_{Mj} + f_{n-1}\,x_{M,j+1} + \sum_{i=j+1}^{n-1} (f_{n-1} + f_{n-i})\,x_{M,i+1}, \quad [26]$$

where

$$\mathbf{G}^{-1} = \begin{bmatrix} 1 & f_{n-1} & f_{n-1}+f_{n-2} & f_{n-1}+f_{n-3} & \cdots \\ 0 & 1 & f_{n-1} & f_{n-1}+f_{n-2} & \cdots \\ 0 & 0 & 1 & f_{n-1} & \cdots \\ & & & \vdots & \\ 0 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & \cdots \end{bmatrix}$$
$$\begin{bmatrix} f_{n-1}+f_3 & f_{n-1}+f_2 & f_{n-1}+f_1 \\ f_{n-1}+f_4 & f_{n-1}+f_3 & f_{n-1}+f_2 \\ f_{n-1}+f_5 & f_{n-1}+f_4 & f_{n-1}+f_3 \\ & & \\ 1 & f_{n-1} & f_{n-1}+f_{n-2} \\ 0 & 1 & f_{n-1} \\ 0 & 0 & 1 \end{bmatrix} \quad [27]$$

6. Shift Register Combinations

We are interested in the properties of shift registers connected in cascade or in parallel. Some of these properties are found simply by inspection. For example, the zero initial state transfer function of

two shift registers in cascade is the product of the individual zero initial state transfer functions. The transfer function of two shift registers connected in parallel (i.e., their outputs are added modulo 2) is the sum of the individual transfer functions. Other properties, including the effects of nonzero initial states, are not so easy to find.

When shift registers are combined, with or without zero initial states, it is often of interest to determine whether there is a single shift register that is equivalent to the combination. We shall be able to find such equivalents for arbitrary initial states. More importantly we shall be able to find the initial states of the component shift registers to yield a prescribed initial state of the equivalent single register. For example, suppose we wish to combine two shift registers of n_1 and n_2 states, respectively, to be equivalent to a single shift register of $n_1 + n_2$ stages. We desire the output sequence of the combination, with zero input, to start with $n_1 + n_2 - 1$ zeros followed by a one. What should be the initial states of the component shift registers?

7. Simple Shift Register Generators in Parallel

We consider the autonomous behavior of two simple shift registers whose outputs are added, as shown in Fig. 3 for two SSRG's. If the individual outputs have transforms $Y^{(1)}(z)$ and $Y^{(2)}(z)$, the total output has a transform $Y(z)$ given by the sum of $Y^{(1)}(z)$ and $Y^{(2)}(z)$. Let the two SSRG's have n_1 and n_2 stages, respectively, and characteristic polynomials $f^{(1)}(z)$, $f^{(2)}(z)$. Then, using Eq. [17] the output transform is

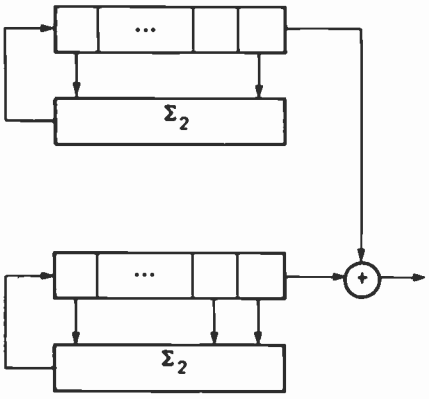


Fig. 3—Two SSRGs in parallel.

$$Y(z) = \frac{g^{(1)}(z) \sum_{j=1}^{n_1} z^j \sum_{i=j}^{n_1} x_i^{(1)}(0) f_{n_1-i+j}^{(1)} + g^{(2)}(z) \sum_{j=1}^{n_2} z^j \sum_{i=j}^{n_2} x_i^{(2)}(0) f_{n_2-i+j}^{(2)}}{\mathrm{lcm}(f^{(1)}, f^{(2)})} \quad [28]$$

where

$$\mathrm{lcm}(f^{(1)}, f^{(2)}) = \text{least common multiple of } f^{(1)}(z), f^{(2)}(z) \quad [29]$$

$$g^{(1)}(z) = \frac{\mathrm{lcm}(f^{(1)}, f^{(2)})}{f^{(1)}(z)} \quad [30]$$

$$g^{(2)}(z) = \frac{\mathrm{lcm}(f^{(1)}, f^{(2)})}{f^{(2)}(z)} \quad [31]$$

With this formula we can find the equivalent single shift register that will give the same output as the parallel combination of Fig. 3, when these have given initial states. What we seek from Eq. [28] is the initial state of the equivalent single shift register.

The necessary formula is derived by noting that for a single shift register, the output transform is given by Eq. [17]. By equating that equation to Eq. [28] using

$$f(z) = \mathrm{lcm}(f^{(1)}, f^{(2)}), \quad [32]$$

we get, with n = degree of $f(z)$,

$$\sum_{j=1}^{n} z^j \sum_{i=j}^{n} x_i f_{n-i+j} = g^{(1)}(z) \sum_{j=1}^{n_1} z^j \sum_{i=j}^{n_1} x_i^{(1)} f_{n_1-i+j}^{(1)} + g^{(2)}(z) \sum_{j=1}^{n_2} z^j \sum_{i=j}^{n_2} x_i^{(2)} f_{n_2-i+j}^{(2)}. \quad [33]$$

The left hand side refers to the single equivalent shift register. The argument has been omitted from each x_i, $x_i^{(1)}$, and $x_i^{(2)}$ because the result is implied to refer to an initial state which is referred to an arbitrary origin.

Eq. [33] is used by equating like powers of z on both sides of the equation.

8. Modular Shift Register Generators in Parallel

Results similar to those in the previous section may be obtained for MSRG's. Starting with Eq. [18], we get the formula corresponding to Eq. [33]:

$$\sum_{j=1}^{n} x_j z^j = g^{(1)}(z) \sum_{j=1}^{n_1} x_j^{(1)} z^j + g^{(2)}(z) \sum_{j=1}^{n_2} x_j^{(2)} z^j, \qquad [34]$$

where $g^{(1)}(z)$ and $g^{(2)}(z)$ are defined by Eqs. [30] and [31], and

$$f(z) = \text{lcm}(f^{(1)}, f^{(2)}). \qquad [35]$$

Again, like powers of z are equated to get a set of linear equations in the desired values for the initial states $x_j^{(1)}$ and $x_j^{(2)}$.

9. Simple Shift Register Generators In Cascade

With two shift registers in cascade, the output of the first is the input to the second. The zero state behavior of the cascade is determined simply by multiplying the individual transfer functions. This product is then multiplied by the z-transform of the input to the cascade to get the z-transform of the output. This behavior presents no particular difficulty, so we turn to the zero input behavior. That is, we consider the autonomous behavior of the cascade with specific initial states in the two shift registers. In particular, we consider the required initial states to give the same output as a single shift register having a prescribed initial state.

We approach the problem by noting that the output of the first shift register is the input to the second. The output of the second, therefore, has two components: one is due to the output of the first, and the other is due to the initial state of the second. Accordingly, we write

$$Y^{(2)}(z) = Y^{(1)}(z)H^{(2)}(z) + \frac{\sum_{j=1}^{n_2} z^j \sum_{i=j}^{n_2} x_i^{(2)}(0) f_{n_2-i+j}^{(2)}}{f^{(2)}(z)}, \qquad [36]$$

where the superscripts refer to the particular shift register. Now, from Eqs. [7] and [17]

$$H^{(2)}(z) = \frac{1}{f^{(2)}(z)} \qquad [37]$$

$$Y^{(1)}(z) = \frac{\sum_{j=1}^{n_1} z^j \sum_{i=j}^{n_1} x_i^{(1)}(0) f_{n_1-i+j}^{(1)}}{f^{(1)}(z)}, \qquad [38]$$

so that the transform $Y(z)$ of the output of the cascade is

$$Y(z) = \frac{\sum_{j=1}^{n_1} z^j \sum_{i=j}^{n_1} x_i^{(1)}(0) f_{n_1-i+j}^{(1)} + f^{(1)}(z) \sum_{j=1}^{n_2} z^j \sum_{i=j}^{n_2} x_i^{(2)}(0) f_{n_2-i+j}^{(2)}}{f^{(1)}(z)\, f^{(2)}(z)}. \quad [39]$$

We are interested now in determining the characteristic polynomial and initial state of the equivalent single SSRG. First of all, the characteristic polynomial $f(z)$ is

$$f(z) = f^{(1)}(z)\, f^{(2)}(z). \quad [40]$$

This shows that the degree n of the equivalent single SSRG is

$$n = n_1 + n_2. \quad [41]$$

The initial state of the equivalent SSRG is related to the initial states of the individual SSRG's in cascade through the equation:

$$\sum_{j=1}^{n_1+n_2} z^j \sum_{i=j}^{n} x_i(0) f_{n-i+j} = \sum_{j=1}^{n_1} z^j \sum_{i=j}^{n_1} x_i^{(1)}(0) f_{n_1-i+j}^{(1)} + f^{(1)}(z) \sum_{j=1}^{n_2} z^j \sum_{i=j}^{n_2} x_i^{(2)}(0) f_{n_2-i+j}^{(2)}, \quad [42]$$

where $n = n_1 + n_2$. Further simplification of Eq. [42] does not appear warranted. It is used by equating the coefficients of like powers of z, yielding a set of equations for the $x_i(0)$ in terms of the $x_i^{(1)}(0)$ and $x_i^{(2)}(0)$.

It is interesting to note that the last n_2 initial values of the single equivalent SSRG are the same as the n_2 initial values of the second SSRG, i.e.,

$$\begin{aligned} x_n(0) &= x_{n_2}^{(2)}(0) \\ x_{n-1}(0) &= x_{n_2-1}^{(2)}(0) \\ &\vdots \\ x_{n-n_2+2}(0) = x_{n_1+2}(0) &= x_2^{(2)}(0) \\ x_{n-n_2+1}(0) = x_{n_1+1}(0) &= x_1^{(2)}(0) \end{aligned} \quad [43]$$

A closely related problem is this. Given:
(1) A characteristic polynomial $f(z)$ that is the product of two factors:

$$f(z) = f^{(1)}(z)\, f^{(2)}(z). \quad [44]$$

The degree of $f(z)$ is n, the degree of $f^{(1)}(z)$ is n_1, and the degree $f^{(2)}(z)$ is n_2.

(2) $n = n_1 + n_2$ successive output digits.

We wish to find the equivalent decomposition into two cascaded SSRG's and find their initial states.

As before, we let $f^{(2)}(z)$ be the characteristic polynomial of the second SSRG. The last n_2 values of the $x_i(0)$ are known and these are the values of the initial state of the second SSRG, as shown by Eq. [43]. The job remaining is to find the $x_i^{(1)}(0)$, $i = 1, 2, \ldots, n_1$.

Fig. 4 indicates the decomposition into two cascaded SSRG's. The output of the second SSRG is the output of the cascade or of the single SSRG equivalent:

$$y(k) = y^{(2)}(k). \qquad [45]$$

The input to the second SSRG is the modulo 2 sum of the output $y^{(1)}(k)$ of the first SSRG and the result $w(k)$ of the feedback in the second SSRG. The input to the second SSRG is simply the value in its first stage at the next shift, $x_1^{(2)}(k + 1)$. Thus,

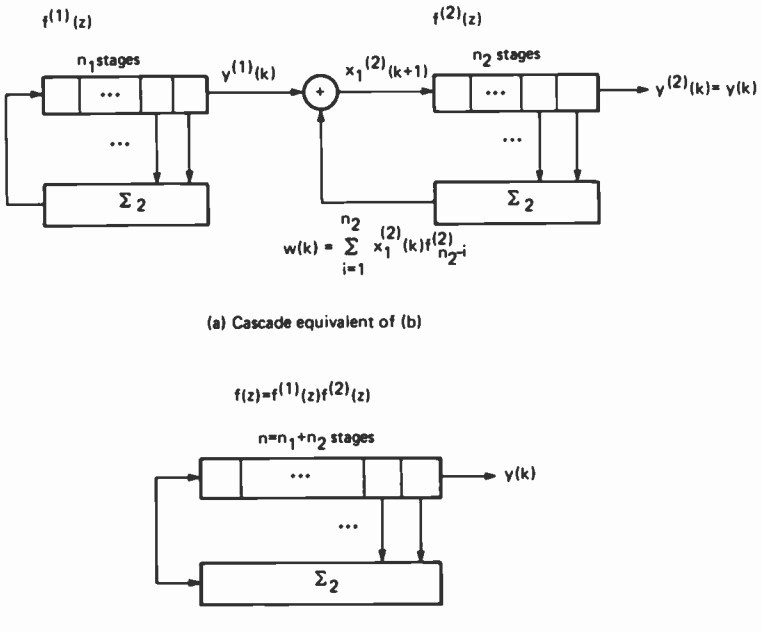


Fig. 4—Decomposition of an SSRG into a cascade equivalent.

$$x_1^{(2)}(k+1) = y^{(1)}(k) + w(k)$$

$$y^{(1)}(k) = x_1^{(2)}(k+1) + w(k) = x_1^{(2)}(k+1) + \sum_{i=1}^{n_2} x_i^{(2)}(k) f_{n_2-i}^{(2)}. \quad [46]$$

Now,

$$y^{(1)}(n_1 - j) = x_j^{(1)}(0)$$
$$x_1^{(2)}(k+1) = y^{(2)}(k+n_2) = x_{n_1}(k) = x_j(k - n_1 + j), \quad [47]$$

and

$$x_i^{(2)}(n_1 - j) = x_{i+j}(0)$$

$$x_j^{(1)}(0) = \sum_{i=0}^{n_2} x_{i+j}(0)\, f_{n_2-i}^{(2)}. \quad [48]$$

This is the formula we sought.

Sometimes it is desirable to get the output of the first register in the equivalent cascade as a physical sequence so that the state of the first register in the equivalent cascade can be operated upon. That is, given $y(k)$, produce $y^{(1)}(k)$ physically. The circuit for doing so can be derived from Eq. [46]. We note that

$$x_i^{(2)}(k) = y(k + n_2 - i). \quad [49]$$

Therefore,

$$y^{(1)}(k) = y(k+n_2) + \sum_{i=1}^{n_2} y(k + n_2 - i) f_{n_2-i}^{(2)}. \quad [50]$$

Then

$$y^{(1)}(k - n_2) = y(k) + \sum_{i=1}^{n_2} y(k - i) f_{n_2-i}^{(2)}$$

$$= \sum_{i=0}^{n_2} y(k - i) f_{n_2-i}^{(2)}, \quad [51]$$

since $f_{n_2}^{(2)} = 1$. This is the desired formula for extracting the component sequences. A circuit to realize Eq. [51] is shown in Fig. 5.

10. Modular Shift Register Generators in Cascade

Here we use a formula like Eq. [36] except that we use, instead of the second term on the right of Eqs. [36] and [38], the corresponding

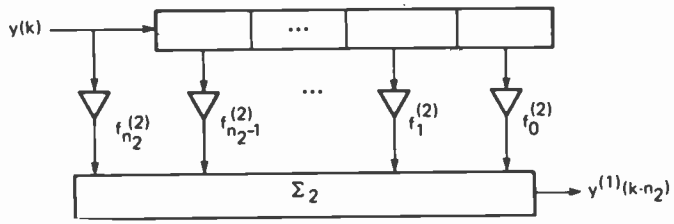


Fig. 5—Extraction of the component sequences from an equivalent cascade.

formula for MSRG's. The formula is, using Eq. [18] and some algebra,

$$Y(z) = Y^{(2)}(z) = \frac{\sum_{j=1}^{n_1} x_j^{(1)}(0)z^j + f^{(1)}(z) \sum_{j=1}^{n_2} x_j^{(2)}(0)z^j}{f^{(1)}(z)\, f^{(2)}(z)}. \qquad [52]$$

As above, the characteristic polynomial of the equivalent single MSRG is $f^{(1)}(z)\, f^{(2)}(z)$ and the degree n of the equivalent is $n_1 + n_2$. The initial state of the equivalent single MSRG is related to the initial states of the individual MSRG's in cascade through the following equation:

$$\sum_{j=1}^{n} x_j(0)z^j = \sum_{j=1}^{n_1} x_j^{(1)}(0)z^j + f^{(1)}(z) \sum_{j=1}^{n_2} x_j^{(2)}(0)z^j. \qquad [53]$$

By equating the coefficients of like powers of z on both sides of Eq. [53], we get the following expression for the values of the initial state vector of the equivalent single MSRG:

$$x_j(0) = x_j^{(1)}(0) + \sum_{i=0}^{j-1} f_j^{(1)}\, x_{j-i}^{(2)}(0), j = 1, 2, \ldots, n. \qquad [54]$$

When this formula is used, it must be remembered that

$$\begin{aligned} x_j^{(1)}(0) &= 0, j > n_1 \\ f_j^{(1)} &= 0, j > n_1 \\ x_{j-i}^{(2)}(0) &= 0, j > n_2 + i \end{aligned} \qquad [55]$$

11. Complements of Sequences

The z-transform is particularly helpful in the investigation of the properties of the complements of sequences. A simple complement

merely reverses the value of all the digits of a sequence. The alternating complement of a sequence is obtained by changing the value of every other digit. An interesting application of the alternating complement occurs in dual-state modulation of a communication signal. More general complements are possible and they are considered here, but the author knows of no application that has yet been made of these generalized complements.

Suppose we have a sequence $y^{(1)}(k)$, $k \geq 0$, and we want to generate its complement $y(k)$; that is, we want to generate a sequence in which zeros have been changed to ones and ones have been changed to zeros. This means that

$$y(k) = 1 + y^{(1)}(k), \; k \geq 0 \text{ (mod 2, of course).} \qquad [56]$$

To obtain generators of the complement, it is more convenient to work with the z-transform of $y(k)$. The z-transform of both sides of Eq. [56] yields

$$Y(z) = \sum_{k=0}^{\infty} z^{-k} + Y^{(1)}(z)$$

$$= \frac{z}{1+z} + Y^{(1)}(z). \qquad [57]$$

Several realizations of Eq. [57] are possible. One involves two shift registers in parallel, one corresponding to $Y^{(1)}(z)$ and the other corresponding to $z/(1+z)$, which we may label $Y^{(2)}(z)$:

$$Y^{(2)}(z) = \frac{z}{1+z}. \qquad [58]$$

When this expression is compared with either Eqs. [17] or [18], it will be seen that Eq. [58] represents the output of a single stage feedback shift register whose characteristic polynomial is

$$f^{(2)}(z) = z + 1 \qquad [59]$$

and whose initial state is unity. The output of the combination is the complement of $y^{(1)}(k)$.

The parallel realization is an obvious one based on the appearance of Eq. [57]. Is there a cascade realization? Is there an equivalent single shift register to give the same output? Let us try, at first, to answer the cascade question using an SSRG.

The situation we want to consider is shown in Fig. 6. The n stage SSRG shown in part (a), with characteristic polynomial $f(z)$ and initial state $x(0)$, produces a certain sequence $y(k)$. The arrangement

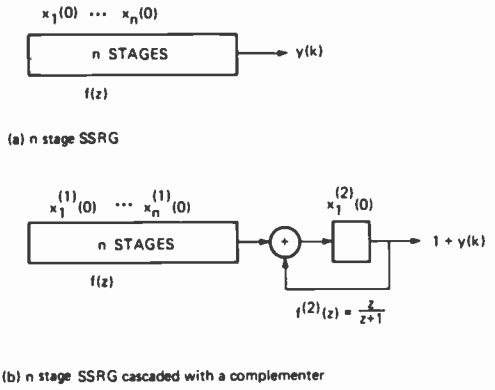


Fig. 6—A cascade arrangement for obtaining the complement of a sequence.

in Fig. 6(b) is to produce the complement $1 + y(k)$ of the sequence $y(k)$ by cascading an n stage SSRG having the same characteristic polynomial with a single stage shift register acting as a complementer. Of course, we do not expect the initial state $\mathbf{x}^{(1)}(0)$ in Fig. 6(b) to be the same as the initial state $\mathbf{x}(0)$ of Fig. 6(a).

What has to be done now is to solve for $\mathbf{x}^{(1)}(0)$ and for the initial state $x_1^{(2)}(0)$ of the complementer. We find that

$$x_1^{(2)}(0) = 1 + x_n(0) = 1 + y(0),$$

and

$$x_{n-k}^{(1)}(0) = x_{n-k}(0) + x_{n-k-1}(0),\ k = 0, 1, 2, \ldots, n-2$$

$$x_1^{(1)}(0) = x_1(0) + \sum_{j=1}^{n} x_j(0) f_{n-j}. \qquad [60]$$

Similar formulas may be found for the arrangement where the complementer precedes the other shift register.

It is possible to find a single shift register equivalent to either the parallel arrangement or the cascade arrangement discussed above. If $f(z)$ is the characteristic polynomial of the uncomplemented sequence, the complemented sequence will have a characteristic polynomial given by $\text{lcm}[f(z), z + 1]$. The initial state of the single equivalent SSRG can be obtained by equating like powers of z in the expressions for the output sequence in terms of the initial states.

12. Alternating Complements

Suppose we have a sequence $y^{(1)}(k)$, $k \geq 0$, and we want to generate a sequence $y(k)$ that is the alternating complement of the sequence $y^{(1)}(k)$; that is, every other digit is complemented. This means that

$$y(k) = y^{(1)}(k) + y^{(2)}(k), \tag{61}$$

where

$$\begin{aligned} y^{(2)}(k) &= 0, \; k \text{ even} \\ &= 1, \; k \text{ odd.} \end{aligned} \tag{62}$$

The z-transform of $y^{(2)}(k)$ is

$$Y^{(2)}(z) = \sum_{k \text{ odd}} z^{-k} = z^{-1} \sum_{k=0}^{\infty} z^{-2k}$$

$$= \frac{z}{1 + z^2}. \tag{63}$$

The z-transform of Eq. [61] becomes

$$Y(z) = Y^{(1)}(z) + \frac{z}{z^2 + 1}. \tag{64}$$

When the second term on the right hand side of Eq. [64] is compared with Eq. [19], we can determine that $Y^{(2)}(z)$ can be realized by a simple two stage feedback shift register, whose characteristic polynomial is

$$f^{(2)}(z) = z^2 + 1, \tag{65}$$

and whose initial state is given by

$$\begin{aligned} x_1^{(2)}(0) &= 1 \\ x_2^{(2)}(0) &= 0. \end{aligned} \tag{66}$$

With this realization of $Y^{(2)}(z)$, the parallel combination of shift register generators to realize $Y(z)$, that is, the sequence that is $y^{(1)}(k)$ with alternate digits complemented, is shown in Fig. 7.

We now look for a single SSRG that is equivalent to the parallel combination shown in Fig. 7. We can see that the single equivalent will have a characteristic polynomial given by $\text{lcm}(f^{(1)}, z^2 + 1)$. The state of the single equivalent can be determined from the states of the parallel arrangement of Fig. 7 by using Eq. [33].

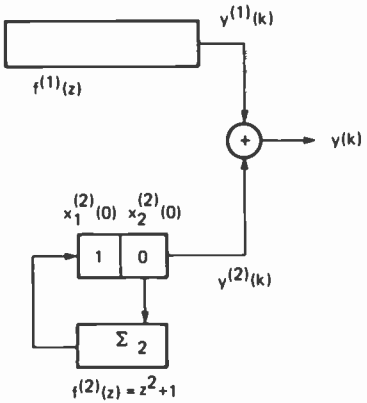


Fig. 7—Obtaining the alternating complement of the sequence $y^{(1)}(k)$, i.e., the alternate digits of $y^{(1)}(k)$ have been reversed.

13. An Application of z-Transforms—Dual State Modulation

In this section, we consider the modulation of a binary sequence upon a carrier wave such that the state (e.g., phase or frequency) of the carrier may or may not change in accordance with the binary sequence message. For example, consider a frequency difference keyed (FDK) waveform in which the current frequency is changed if the next symbol is a zero and is left unchanged if the next symbol is a one. Fig. 8 illustrates the process of modulation.

There is a base frequency F_1. At the start of modulation, the frequency changes to its other value F_2 because the first symbol in the modulating sequence is a zero. If we let

$y(k)$ = message sequence = modulating sequence
$m(k)$ = modulated sequence = transmitted sequence

we can see that

$$m(k) + m(k-1) = 1 + y(k). \qquad [67]$$

The received sequence, after converting frequency to symbols, is $m(k)$, where either frequency may be demodulated as a one. Recovery of the message $y(k)$ is accomplished using Eq. [67]:

$$y(k) = 1 + m(k) + m(k-1). \qquad [68]$$

An algorithm for recovery of the message is shown in Fig. 9, which simply provides the operation of Eq. [68].

Additional insight is obtained by using z-transforms. Let

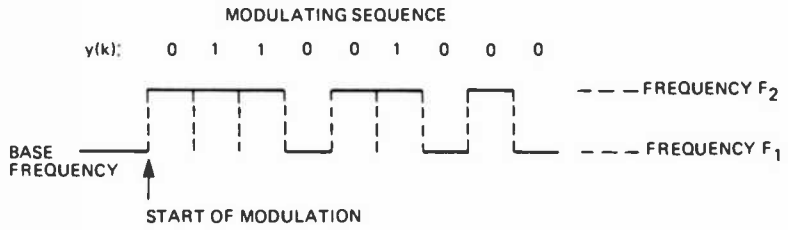


Fig. 8—Frequency difference keying (FDK) by a binary modulating sequence.

$Y(z)$ = z-transform of $y(k)$
$M(z)$ = z-transform of $m(k)$.

The z-transform of Eq. [68] yields

$$Y(z) = \frac{z}{z + 1} + M(z)\,(1 + z^{-1}). \qquad [69]$$

Solving for $M(z)$ yields

$$M(z) = \frac{zY(z)}{z + 1} + \frac{z^2}{z^2 + 1}. \qquad [70]$$

Now, $Y(z)$ is the transform of a shift register sequence with a char-

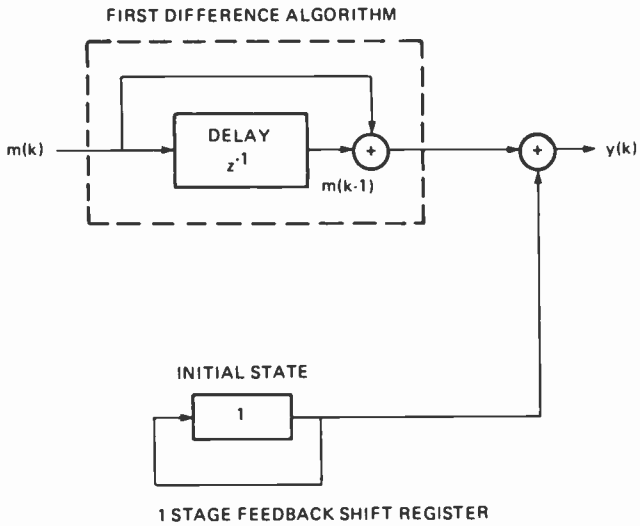


Fig. 9—Recovery of message $y(k)$ from an FDK signal. The first difference algorithm may be accomplished by a one-stage shift register with initial state zero.

acteristic polynomial $f(z)$, say. This means that $Y(z)$ can be written as

$$Y(z) = \frac{N(z)}{f(z)}, \qquad [71]$$

where $N(z)$ is a polynomial in z. Inserting Eq. [71] in Eq. [70], we obtain

$$M(z) = \frac{z(z + 1)N(z) + z^2 f(z)}{(z^2 + 1)f(z)}, \qquad [72]$$

presuming that $z + 1$ does not divide $f(z)$. Examination of Eq. [72] shows that $m(k)$, the sequence whose transform is $M(z)$, has a characteristic polynomial $(z^2 + 1)f(z)$. This has degree two greater than the degree of $f(z)$. Thus, we may say that the process of frequency difference keying increases the degree of a sequence by two.

Now consider what happens after the first difference process. The transfer function is $1 + 1/z$, so that with $m(k)$ as its input, the output transform is

$$\frac{M(z)(z + 1)}{z} = \frac{(z + 1)N(z) + zf(z)}{(z + 1)f(z)}. \qquad [73]$$

The resulting sequence has $(z + 1)f(z)$ as its characteristic polynomial. Thus, the first difference process yields a sequence that has degree one greater than the original modulating sequence. The process of complementing will add the z-function $z/(z + 1)$ to Eq. [73]. The result is $N(z)/f(z)$ which is $Y(z)$.

14. Generalized Complements

The idea of complementing a sequence can be generalized and the use of z-transforms with such generalizations shows some interesting algebraic properties. However, the author has so far found no applications beyond those for the alternating complement discussed earlier. We start with the following definition: An nth order, single-digit complement is one of the sequences obtained by complementing every nth digit of a given sequence. The ordinary complement is the first-order, single-digit complement. The alternating complements (there are two) are second-order, single-digit complements. For a given sequence, there are m possible mth order single-digit complements.

We shall consider single-digit complements first; this will make

simpler the understanding of multiple-digit complements. We define an mth order single-digit *complementing* sequence as follows:

$$\begin{aligned} y(k;\, m,\, 1;\, j) &= 1,\ k = j \bmod m \\ &= 0, \text{ otherwise.} \end{aligned} \qquad [74]$$

Here, k is the integer time index; m is the order of the complementing sequence; $j = 0, 1, \ldots, m - 1$ and j indicates the particular one of the m possible complementing sequences; the number following m (a 1 in this case) indicates the number of digits complemented (i.e. reversed) in each group of m.

The complementing sequence is added to a given complemented sequence $y^{(1)}(k)$ to yield a *complementary* sequence (i.e., an nth order single-digit complementary sequence). With $y(k)$ indicating the resulting sequence, we have

$$y(k) = y^{(1)}(k) + y(k;\, m,\, 1;\, j). \qquad [75]$$

An example will help to clarify the notation. Let $m = 3$; then j has the possible values 0, 1, 2. Then, we have

$$\begin{aligned} y(k;\, 3,\, 1;\, 0) &= 1,\ k = 0 \bmod 3 \\ &= 0, \text{ otherwise} \\ y(k;\, 3,\, 1;\, 1) &= 1,\ k = 1 \bmod 3 \\ &= 0, \text{ otherwise} \\ y(k;\, 3,\, 1;\, 2) &= 1,\ k = 2 \bmod 3 \\ &= 0, \text{ otherwise} \end{aligned}$$

Now we apply the definition Eq. [3] of the z-transform to Eq. [74]. We find that nonzero values occur only for $k = j + rm$, where $r = 0, 1, \ldots$. Then the z-transform of $y(k;\, m,\, 1;\, j)$ is

$$Y(z;\, m,\, 1;\, j) = \sum_{r=0}^{\infty} z^{-j-rm} = \frac{z^{m-j}}{1 + z^m}. \qquad [76]$$

Now, the z-transform of $y^{(1)}(k)$, it will be recalled, has the form

$$Y^{(1)}(z) = g^{(1)}(z)/f^{(1)}(z), \qquad [77]$$

where $f^{(1)}(z)$ is the characteristic polynomial of $y^{(1)}(k)$ and the polynomial $g^{(1)}(z)$ represents a set of initial conditions and determines which of the possible sequences is under discussion with characteristic polynomial $f^{(1)}(z)$. A comparison with Eq. [17] shows that $g^{(1)}(z)$ corresponds to the initial state of an SSRG which generates $y^{(1)}(k)$. A comparison with Eq. [18] shows that the coefficients of $g^{(1)}(z)$ are, directly, the elements of the initial state of the MSRG

which generates $y^{(1)}(k)$. We now use Eqs. [76] and [77] to get $Y(z)$, the z-transform of $y(k)$:

$$\begin{aligned} Y(z) &= Y^{(1)}(z) + Y(z;\, m,\, 1;\, j) \\ &= g^{(1)}(z)/f^{(1)}(z) + z^{m-j}/(1 + z^m) \\ &= g(z)/\mathrm{lcm}(f^{(1)}(z),\, 1 + z^m) \end{aligned} \qquad [78]$$

Here, $g(z)$ is a polynomial of degree not exceeding m plus the degree of $f^{(1)}(z)$. Under the presumption that the numerator and denominator of Eq. [78] have no common factors, we can recognize that the denominator represents the characteristic polynomial $f(z)$ of a shift register sequence:

$$f(z) = \mathrm{lcm}(f^{(1)}(z),\, 1 + z^m). \qquad [79]$$

The numerator $g(z)$ represents the initial conditions of the single shift register generator that would generate $Y(z)$. By equating [78] with Eq. [17], $g(z)$ would give the initial state of the SSRG and $f(z)$ would give the characteristic polynomial for generating the result $y(k)$. By equating [78] with Eq. [18], the corresponding initial state of the equivalent MSRG would be obtained.

Of course, a direct way of generating $y(k)$ is to add the outputs of the shift register that generates $y^{(1)}(k)$ and the shift register that generates $y(k;\, m,\, 1;\, j)$. The complementing SSRG would have its initial state consist of a single nonzero digit in the position represented by the exponent $m - j$ in Eq. [76]. Its characteristic polynomial is $1 + z^m$; this represents a circulating shift register whose output is fed back directly to the first stage.

If it is desired to recover the original sequence $y^{(1)}(k)$ from $y(k)$ when $y(k)$ is known to be an mth order single-digit complementary sequence, we simply write, from Eq. [75],

$$y^{(1)}(k) = y(k) + y(k;\, m,\, 1;\, j). \qquad [80]$$

Because j is unknown, there will be m possible second terms on the right hand side of Eq. [80]. All of these are to be tried, yielding m different sequences. To determine the correct one, each is tested to see that it produces the proper recursion. That is, Eq. [1] or [2] is to be satisfied.

We can generalize the above by considering q digit, mth order complementing sequences and using the symbol $y(k;\, m,\, q;\, j)$. The symbols q and m indicate that q digits out of each block of m have been reversed. The number j indexes the various possible patterns of r positions in each group of m. This number is $\binom{m}{r}$, i.e., $j = 0, 1, \ldots, \binom{m}{r} - 1$. In $y(k;\, m,\, q;\, j)$, nonzero values occur only for certain values of k, say k_j in each block of m. That is,

$$y(k;\, m,\, q;\, j) = 1,\ k = k_j + rm$$
$$= 0, \text{ otherwise,} \qquad [81]$$

where r is an integer and $j = 0, 1, 2, \ldots, q - 1$. This shows that in each block of m digits, there are just q positions that have nonzero digits. Of course Eq. [81] does not indicate which of the $\binom{m}{q}$ possible patterns of nonzero digits is meant.

The z-transform of Eq. [81] is

$$Y(z;\, m,q;\, j) = \sum_{j=1}^{q} \sum_{r=0}^{\infty} z^{-k_j - rm}$$

$$= \frac{1}{1 + z^m} \sum_{j=1}^{q} z^{m-k_j}. \qquad [82]$$

As before, a comparison with Eq. [17] shows that $y(k;\, m,\, q;\, j)$ may be obtained by a circulating shift register whose initial state is described by the polynomial in the numerator of Eq. [82]. If an SSRG is used, then Eq. [17] applies; if an MSRG is used, Eq. [18] applies.

To create a complementary sequence $y(k)$ from a given sequence $y^{(1)}(k)$, $y(k;m,q;j)$ is added:

$$y(k) = y^{(1)}(k) + y(k;m,q;j), \qquad [83]$$

just as in Eq. [75] for single-digit complements. The resulting sequence will have a z-transform like Eq. [78] and the characteristic polynomial will be given by Eq. [79]. The remarks following Eq. [79] apply here as well.

15. Conclusions

In this paper, state variables and z-transforms have been used to obtain several properties of shift-register sequence generators. Among these results are equivalences between simple shift register generators (SSRG's) and multiple return or *modular* shift register generators (MSRG's), between parallel or cascade combinations and a single larger register, and some equivalences involving complemented sequences, alternating complements, and generalized complements. The results are all given in terms of mathematical formulas that enable one, in principle, to work out any specific case. For long shift registers, the hand methods can be quite tedious and, in fact, may not be suitable at all. Machine conversion from an SSRG state to the equivalent MSRG state is straightforward, because the conversion formula is a convolution of the SSRG state

with the characteristic polynomial. Another conversion involving a simple convolution is the extraction of the component sequences from a sequence that has been generated by a cascade of two shift register generators.

All of the other conversions and equivalences involve operations that can be performed by linear binary shift registers. These operations may include binary polynomial arithmetic which Peterson[3,4] shows may be performed with shift registers.

Appendix A—Derivation of the z-Transforms of the Autonomous Outputs

Eqs. [17] and [18] are the z-transforms of the autonomous output sequences of the SSRG and MSRG, respectively, when the initial states are respectively x_{si} and x_{Mi}, $i = 1, 2, \ldots, n$. We start with the SSRG. Eq. [17] may be derived from the last of Eq. [4]. In Eq. [4], the set $y(n - i)$, $1 \leq i \leq n$, can be considered as the initial state for the sequence $y(k)$. If this is so, we may simply set

$$x_{si}(0) = y(n - i) \qquad [84]$$

This yields Eq. [17] from Eq. [4].

Obtaining Eq. [18] is a little more involved. The demonstration is made by showing that $Y(z)$ as obtained from Eq. [14] is the same as Eq. [18] for the MSRG. Starting with Eq. [10a], we take the z-transform of both sides, noting that

$$Z[\mathbf{x}(k + 1)] = z\mathbf{X}(z) - \mathbf{x}(0). \qquad [85]$$

Then Eq. [13] is used to get

$$Z[\Phi(k)] = \boldsymbol{\phi}(z) = z[z\mathbf{I} - \mathbf{A_M}]^{-1}. \qquad [86]$$

It follows from Eq. [14] that

$$Y(z) = zC[z\mathbf{I} - \mathbf{A_M}]^{-1}\mathbf{x}(0). \qquad [87]$$

Eq. [16] is to be used in Eq. [87], but the entire inverse matrix is not needed. As Eq. [11] shows, C is a row vector whose only nonzero component is the last, unity. Thus, only the last row $\phi_{n1}, \phi_{n2}, \ldots, \phi_{nn}$ of the inverse needs to be calculated. When this fact is used (after some labor) and the result put into Eq. [87], the result will be Eqs. [18].

Appendix B.—Glossary of Principal Symbols

$\mathbf{A}$	system matrix
$\mathbf{A_M}$	system matrix of the MSRG

$\mathbf{A}_s$	system matrix of the SSRG
$\mathbf{C}$	output matrix, relating the contribution of a state $\mathbf{x}(k)$ to the output $y(k)$
$f(z)$	characteristic polynomial: coefficients are $f_0, \ldots, f_n$
$\mathbf{G}$	state conversion matrix
$H(z)$	transfer function
$H_M(z)$	transfer function of an MSRG
$H_s(z)$	transfer function of an SSRG
$h(k)$	impulse (sequence) response
$h_M(k)$	impulse response of an MSRG
$h_s(k)$	impulse response of an SSRG
k	integer shift or time index
lcm	least common multiple
MSRG	modular or multiple return shift register generator
n	number of shift register stages. Degree of characteristic polynomial
SSRG	simple shift register generator
$\mathbf{x}(k)$	state vector. Components are $x_i(k)$
$\mathbf{x}_M(k)$	state vector of an MSRG. Components are $x_{Mi}(k)$
$\mathbf{x}_s(k)$	state vector of an SSRG. Components are $x_{si}(k)$
$Y(z)$	z-transform of the output sequence
$Y_M(z)$	z-transform of the output sequence of an MSRG
$Y_s(z)$	z-transform of the output sequence of an SSRG
$y(k)$	output sequence
$y^{(1)}(k)$	a component sequence
$y(k;m,j,r)$	an mth order, r digit, complementing sequence
z	z-transform variable
$\Phi(k)$	state transition matrix ($=\mathbf{A}^k$)
$\boldsymbol{\Phi}(z)$	z-transform of $\Phi(k)$. Elements are $\phi_{ij}(z)$.

References:

[1] D. A. Huffman, "The Synthesis of Linear Sequential Coding Networks," in *Information Theory*, D. Cherry (Ed.), pp. 77–95, Academic Press, N.Y. (1956).
[2] D. A. Huffman, "A Linear Circuit Viewpoint on Error-Correcting Codes," *IRE Trans. Inf. Theory* **IT-2**, p. 20, March 1956.
[3] W. W. Peterson, *Error Correcting Codes*, Chapter 7, MIT Press, Cambridge, Mass. (1961).
[4] W. W. Peterson and E. J. Weldon, Jr., *Error Correcting Codes*, Chapter 7, 2nd Ed., MIT Press, Cambridge, Mass. (1972).
[5] B. Elspas, "The Theory of Autonomous Linear Sequential Networks," *IRE Trans. Circuit Theory*, **CT-6**, No. 1, p. 45, March 1959.
[6] T. G. Birdsall and M. P. Ristenbatt, "Introduction to Linear Shift Register Sequences," Tech. Report 90, Electronic Defense Group, Dept. of EE, Univ. of Michigan Research Instit., Oct. 1958. Available from National Tech. Info. Service (NTIS), Springfield, Va. (Doc. AD225380).
[7] E. R. Berlekamp, *Algebraic Coding Theory*, McGraw-Hill Book Co., N.Y. (1968).
[8] J. A. Cadzow, *Discrete Time Systems*, Chapter 12, Prentice-Hall, Inc., Englewood Cliffs, N.J. (1973).

Image Formation in Contact-Printed PMMA Resist Sublayers

L. K. White and D. Meyerhofer
RCA Laboratories, Princeton, NJ 08540

Abstract—Contact printing of PMMA sublayers is used in portable conformal mask (PCM) multilayer resist techniques. The deep-UV exposure and development of these PMMA sublayers are simulated and compared to experimental results. For thick planarization sublayers (≥ 2 μm), diffraction of the deep-UV radiation produces significant exposure variations across the resist mask apertures. This effect leads to an edge-sharpening phenomena. The Fresnel diffraction approximation is used to calculate exposure variations within the sublayer and a modified SAMPLE program is used to obtain the sublayer image. The effects of dye additions to the sublayer are also analysed. Estimates of the wall profiles and dimensional deviations have been obtained for various isolated space feature sizes.

Introduction

As the requirements for integrated-circuit geometries become more demanding, new lithographic techniques are required. One such technique, the portable conformable mask (PCM) multilayer resist processing scheme, has been used to print fine-line optical lithography patterns (≤ 2 μm) on topographical features.[1–3] A polymethylmethacrylate (PMMA) sublayer is used to planarize the topographical features on the substrate. It is covered with a layer of positive resist that is patterned as usual. This layer acts as a conformal mask for deep-UV (DUV) exposure (200–250 nm) of the PMMA sublayer. Organic developers are usually used to remove the exposed portions of the PMMA sublayer.

Several important processing considerations arise for the practical implementation of this bilayer process. Interfacial mixing of the positive resist with the PMMA can leave positive resist residues

in the PMMA that effectively mask the DUV exposure of the PMMA sublayer.[3,4] The PMMA sublayer is transparent to the actinic radiation that exposes the positive resist. When the radiation is reflected from the substrate, exposure variations are produced in the resist layer due to interference effects (standing wave phenomena) and side-wall reflections.[5–7] Suitable dye additions to the sublayer are required, therefore, for effective linewidth control on reflective substrates. If the positive-resist mask is left intact after exposure on the sublayer, the developer for the PMMA sublayer must not attack the positive-resist material. Chlorobenzene has been the only developer reported that appears to meet this requirement.[3,4,6]

Recently, reports[8,9] of this bilayer process being used in a production mode have appeared. Kodak 809 resist material was used to minimize interfacial mixing and a Coumarin dye[10] was added to the sublayer. The positive resist mask was removed prior to sublayer development so that conventional PMMA developer formulations could be used. Plasma etching conditions for the substrate had to be carefully controlled due to the poor plasma etch resistance of the exposed PMMA sublayer.

These developments, particularly the use of the erodable PMMA sublayer as the dry-etch mask, have placed emphasis on the resolution limits and details of image formation during contact printing. In this work, we analyze the effects of diffraction-induced exposure phenomena and the subsequent development process on the actual contact-printed PMMA sublayer image. Experimentally obtained images are compared to simulated images obtained from exposure and development models. Dimensional deviations as a function of the sublayer thickness and the feature size are presented. An estimate of the effects on these parameters of dye additions to the sublayer is also included.

Experimental

Three-inch-diameter silicon wafers were coated with two layers of KTI PMMA (496 K) 9% solids to produce a 2-μm-thick sublayer. Kodak 809 at 32% solids was used as the DUV conformal mask at a thickness of 1.4 μm. The thick resist layer was used so that near-perfect mask properties could be obtained. A PE220 aligner or an Optimetrix 8010 stepper were used to expose the Kodak 809 resist layer. After development, the sublayer was exposed by a Canon 501 FA aligner with a Xe-Hg (800A) lamp and 250-nm cold mirror. The intensity as measured by an OAI 253-nm probe was 10 mW/cm^2. A dose of 2000 mJ/cm^2 was required to develop the PMMA with a 90-

second dip in chlorobenzene. A toluene rinse and spin dry was used to end the development cycle. The image patterns examined are one-dimensional lines and spaces.

For the dye-addition studies, the Coumarin 6 dye was added directly to the 9% solids KTI solution to produce 2.9% dye by weight. The identical double-coating procedure produced a film thickness that was only 200 Å greater than that of the PMMA coating without the dye addition. The optical absorption spectra of the dyed and undyed PMMA films were determined by spinning the film on transparent sapphire substrates (Fig. 1). At 2.9% dye concentration, a 2-μm-thick dyed layer has an optical density of 0.9 at 440 nm and of 0.27 at 250 nm. A density of 0.9 is sufficient to eliminate most of the reflections off the substrate that affect the resist exposure.[5,11] During the deep UV exposure we did not detect any appreciable change in the optical density at 250 nm. Accordingly, our simulations of dyed sublayer development do not include bleaching.

Modelling of Exposure and Development

To fully understand the PCM system it is necessary to have a model of exposure and development. As a first approximation, we assume that the top photo resist layer is totally opaque to DUV radiation, the only kind that exposes the PMMA. Then the resist–sublayer interface acts a binary mask, either transmitting or blocking the DUV radiation. This truly conformable situation is a good approximation, since most conventional positive resists are very heavily

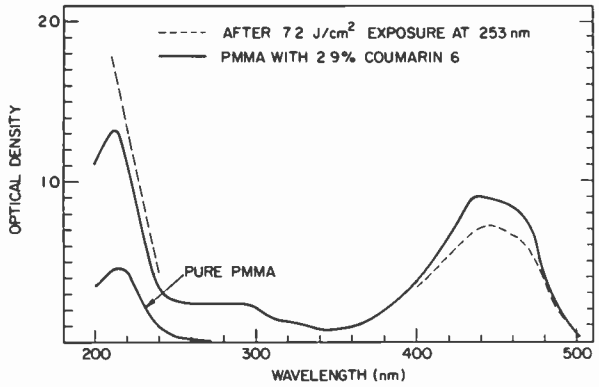


Fig. 1—Absorptions curves of PMMA with and without Coumarin 6 dye. The films were approximately 2.0 μm thick. The dashed line shows the absorption of the dyed film after heavy DUV exposure.

absorbing (10^5 cm^{-1}) in the range $\lambda < 250$ nm, where PMMA is sensitive.

We determine the radiation intensity in the sublayer by making use of the Fresnel-Kirchoff equations.[12] Similar calculations have been used to describe proximity printing of resist films.[13] This approximation is poor at the very top of the sublayer, just below the mask, but provides a good representation of the light distribution for depths more than 0.1 to 0.2 μm below the mask. Examples of calculated diffraction patterns are shown in Fig. 2. The exposure in the absence of a mask would be 0.5 intensity units. The images are long isolated spaces of 2.0-, 1.5-, and 0.75-μm widths and the figure shows a cross-section of these spaces. The radiation is taken to be monochromatic (230 nm) and normally incident, an adequate approximation for the case of this optical system and PMMA. The calculations can readily be extended to multiple wave-lengths and to a range of incident angles, if necessary. In Fig. 2 we have assumed the PMMA is transparent to the radiation and that there is no reflection from the substrate. The actual absorption can readily be included in the calculation.[12] Any reflection from the substrate is added in an incoherent way to complete the radiation pattern.

Fig. 3a shows a three-dimensional display of the radiation pattern from a 1.5-μm-wide space in pure PMMA. The dose at the top of the layer is taken as 1000 mJ/cm^2. The measured value of absorption coefficient of PMMA at 230 nm, 0.27 μm^{-1}, is used in the calculation. The diffraction patterns are more obvious in this representation, and it can be seen how they change with distance from the mask. When the effect of the dye is included by changing the absorption coefficient to 0.66 μm^{-1}, the pattern of Fig. 3b is obtained. The exposure at the substrate interface is considerably reduced.

The radiation pattern and exposure time are used as input to a calculation of development contours. This calculation uses a modified SAMPLE program.[14] To determine the input parameters, we measured the development rates of PMMA films after exposure with a high-pressure Hg lamp, whose main contribution is at 253 nm. The program plots the resist contours that can be expected after various times of development.

Results

Fig. 4 shows SEM micrographs of partially developed sublayer images for various-size isolated spaces. The positive resist conformal mask was patterned with a PE220 aligner. An imprint of the diffraction pattern has been recorded in the sublayer image. Similar

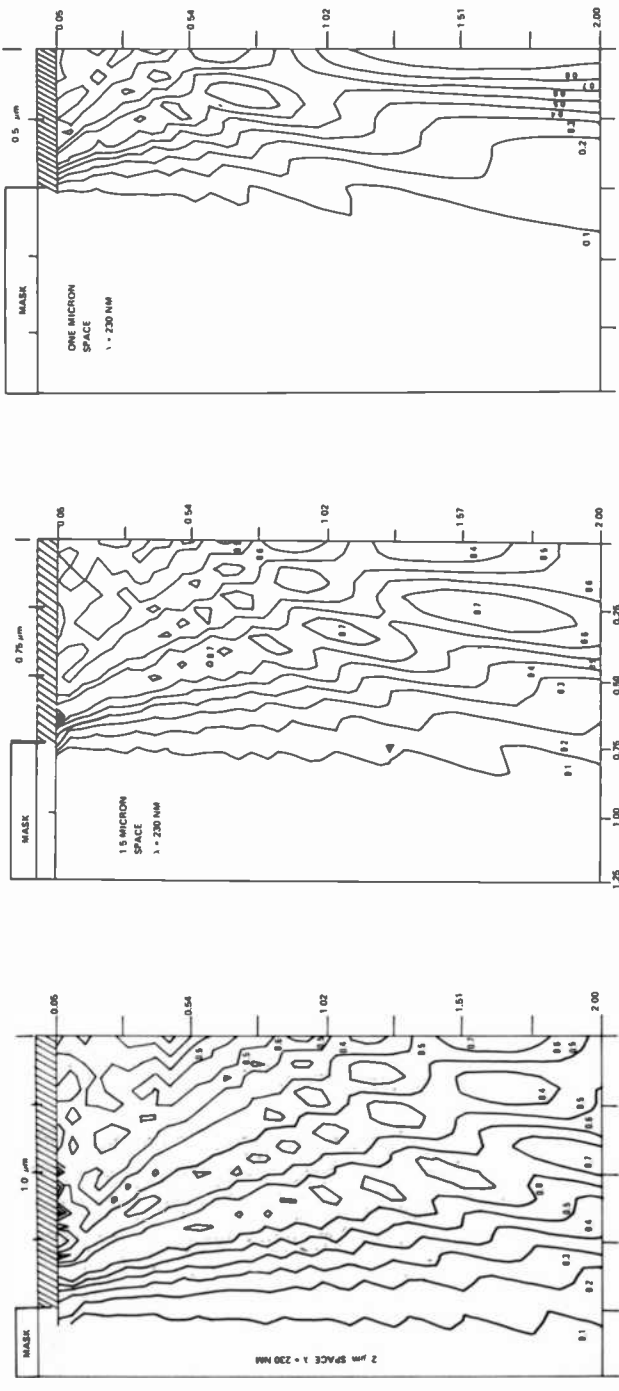


Fig. 2—Calculated diffraction patterns of 230 nm radiation in PMMA without absorption for slits of 1.0-, 1.5-, and 2.0-μm width. The intensity at the film surface is 0.5 units.

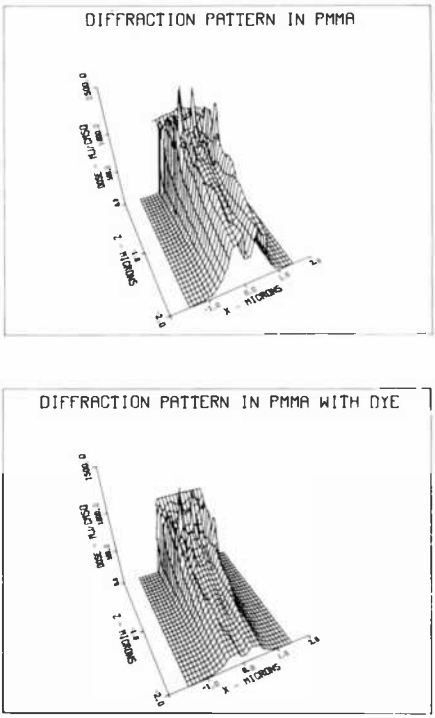


Fig. 3—(a) Three-dimensional diffraction pattern in PMMA from 1.5 μm wide slit. The absorption coefficient of the PMMA at 230 nm is taken to be 0.27 μm^{-1}. The incident dose is 1000 mJ/cm^2. (b) The diffraction pattern for Coumarin doped PMMA, with same parameters.

results have been reported by Lin.[2] For large feature sizes (>2 μm), the PMMA close to the edge of the resist mask develops out first. As the dimensions shrink, the two edges become one in the center of the line and, finally, the edge sharpening disappears altogether. Fig. 5 shows similar results for the positive resist mask patterned with an Optimetrix 8010 stepper. The 2-μm PMMA sublayer is developed out to a thickness of about 1.0 μm.

Fig. 6 shows crossectional simulations of the isolated space feature that correspond to those presented in the experimental results in Figs. 4 and 5. The same development cycle was used for all the simulations. The 2-μm space development simulation results agree reasonably well with the experimental results. Deep trenches are produced near the sidewalls and shallow trenches are observed near the center of the feature. Simulations for the 1.5-μm space and 0.75-μm space do not show as good agreement as the previous results.

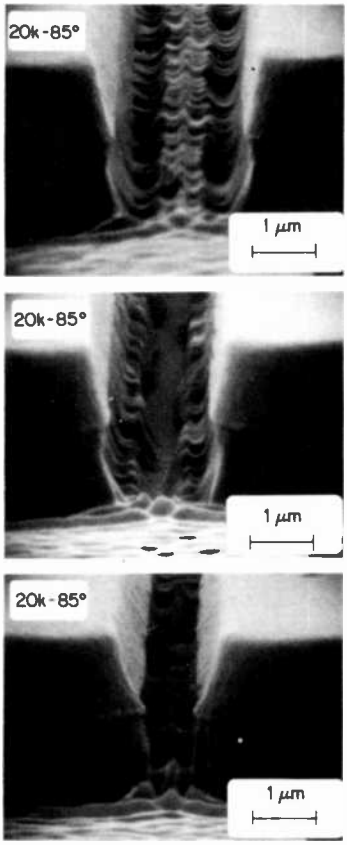


Fig. 4—Partially developed PMMA sublayer with 0.5 μm of initial 2.0-μm thickness remaining.

The combining of the two sidewall trenches for the 1.5-μm space is not observed in the simulation, although the smaller trenches in the center have disappeared. At the 0.75-μm feature size, the simulated images breakthrough to the substrate, while the experimental images are not as deep. A 10% reduction in the development cycle, however, produces no breakthrough to the substrate. This anomaly may indicate that reduced development rates may occur for these small space widths. Although we cannot be certain of the exact reason for these discrepancies, we believe that anomalies in the positive-resist mask are the most probable cause. The 1.5- and 0.75-μm isolated spaces are difficult to print at all on a Perkin Elmer 220 machine. Slight resist residues that may have dissolved during sublayer development could effectively change the actual

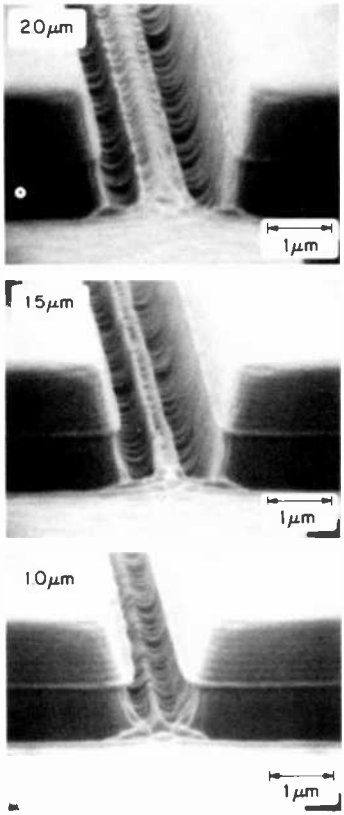


Fig. 5.—Partially developed PMMA sublayer with 1.0 μm of initial 2.0-μm thickness remaining.

dimension of the isolated line. Note that the experimental diffraction pattern gives a result that can be predicted by using a slightly smaller resist-mask dimension. With these considerations in mind we are somewhat assured the modelling procedure employed here gives an useful approximation of the actual contact-printed image of the sublayer.

Fig. 7 shows simulations of fully developed isolated spaces. The development time was adjusted so that the developed width of the 2.0-μm feature had the correct dimension. Table 1 gives the developed sublayer spacewidth dimension at the depths of 2.0, 1.0, and 0.0 μm. The estimated wall profiles at the bottom of the sublayer (between 1.0- and 2.0-μm thickness) and the top of the sublayer (between the 0.0- and 1.0-μm thickness) are shown. At the 1.5-μm resist-mask dimension, the developed sublayer spacewidth still

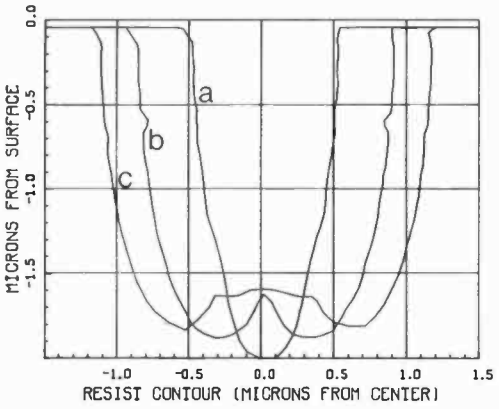


Fig. 6—Calculated resist profiles for partially developed PMMA films. Three different mask openings are shown: (a) 2.0, (b) 1.5, (c) 0.75 μm. The same development times were used in all cases.

matches the resist-mask dimension. Even at the 0.75-μm resist-mask dimension, only a slight skrinkage in spacewidth is observed. The behavior at half thickness (at 1-μm depth) is quite similar. These results suggest that sublayer image degradation does not occur until feature sizes are below 1.5 μm. Even at the 0.75-μm level, image degradation appears to be slight. Note that wall profiles at the bottom of the step are considerably less steep than those at the top of the step.

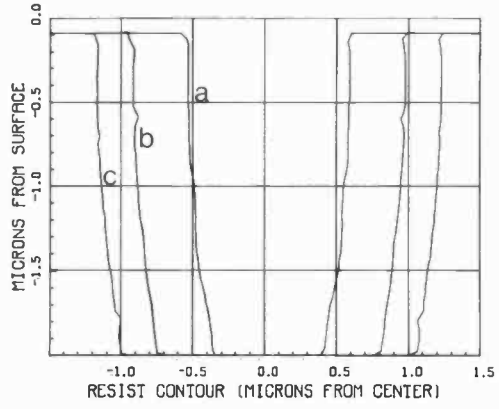


Fig. 7—The same profiles as in Fig. 6, but development is allowed to proceed until the 2.0-μm feature had the nominal width at the substrate. The other two curves are for the same development time.

Table 1—Dimensional Data for Simulated Sublayer Development

Depth in Sublayer (μm)	Developed Spacewidth Dimension		
	2.0 μm Resist Mask	1.5 μm Resist Mask	0.75 μm Resist Mask
0.0 (top)	2.38	1.86	1.10
1.0	2.31	1.80	1.02
2.0 (bottom)	2.00	1.50	0.70
	Estimated Wall Profiles		
1.0–2.0 depth	81°	81°	80°
0.0–1.0 depth	88°	88°	87°

The same development simulations were also performed on PMMA containing dyes. We used the measured value of absorption coefficient, corresponding to the data of Fig. 3b, as well as a sample with an even higher value of the absorption coefficient, 1.03 μm^{-1} (more dye). The three cases that are compared in Fig. 8 are for a 1.5-μm-wide space. The radiation doses have been adjusted to produce the same opening at the substrate interface in all three cases. As expected, the higher absorption requires a longer exposure. In addition, the slope at the substrate decreases and there is more of an undercut relative to the initial positive resist mask. The measured slopes between 1.0- and 2.0-μm sublayer depths for a, b, and c are 81°, 76°, and 72°, respectively.

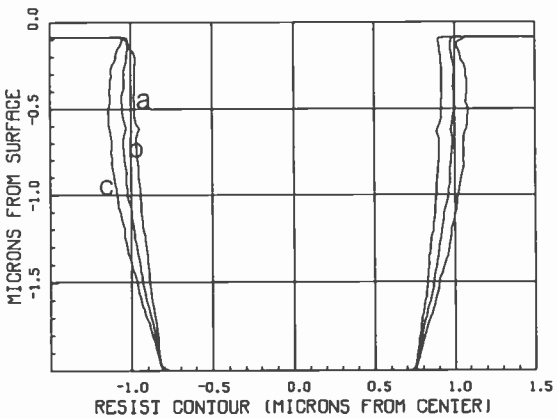


Fig. 8—Comparison of profiles produced in PMMA films with varying absorption coefficients: (1) 0.27, (b) 0.66, (c) 1.03 μm^{-1}. (a) corresponds to pure PMMA and (b) to PMMA with 2.9% Coumarin 6 dye.

Experimental resist profiles of dyed and undyed samples were also compared. For the 2.9% dye the required exposure for the same development time increased by a factor of 2 to 3, which is consistent with the increased development time in the calculation. The profiles showed the correct trend, but the effect is too small to measure quantitatively.

The estimations of resist-wall profile and dimensional control obtained from the model may be adequate for several device designs and etching processes. Improved sublayer parameters can be obtained by overdeveloping or overexposing the sublayer, provided the accompanying dimensional shrinkage can be tolerated. Fig. 9 shows SEM micrographs of fully developed sublayers with and without the dye. The wall profiles are in reasonable agreement with those predicted by the modeling procedure.

Conclusions

The PMMA resist profiles in the PCM structure appear to be well described by diffraction/contact printing and standard development

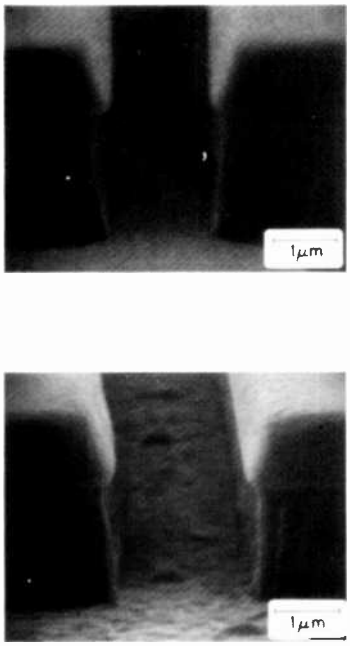


Fig. 9—Fully developed PMMA sublayers. Top is PMMA sublayer without dye and bottom is PMMA sublayer with 2.9% Coumarin 6 dye.

modeling. Penalties in edge acuity (wall-profile steepness) and dimensional deviations on steps due to optical density changes at the actinic wavelength of the sublayer can be predicted with reasonable accuracy.

Acknowledgments

The authors thank Ben Seabury for the SEM analyses, Fred Brehm for the exposure contour graphing, Ron Johansson, Emil Gavalchin, and Nancy Miszkowski for their technical assistance.

References:

[1] B. J. Lin and T. H. P. Chang, "Hybrid E-beam/Deep-UV Exposure Using Portable Conformable Masking (PCM) Technique," *J. Vac. Sci. Technol.*, **16**(6), p. 1669 (1979).
[2] B. J. Lin, "Portable Conformable Mask—A Hybrid Near-Ultraviolet and Deep-Ultraviolet Patterning Technique," *SPIE Proc.*, **174**, p. 114 (1979).
[3] B. J. Lin, E. Bassous, V. W. Chao, and K. E. Petrillo, "Practicing the Novolac Deep-UV Portable Conformable Masking Technique," *J. Vac. Sci. Technol.* **19**, p. 1313 (1981).
[4] M. L. Long, "Practical Deep Ultraviolet—The Multilayer Approach," *SPIE* Proc., **333**, p. 54 (1981).
[5] M. M. O'Toole, E. D. Liu, and M. S. Chang, "Linewidth Control in Projection Lithography Using a Multilayer Resist Process," *IEEE Trans. Electron Dev.*, **ED-28**(11), p. 1405 (1981).
[6] C. H. Ting, I. Avigal, and B. C. Lu, "Multilayer Resist Strategy for High Resolution Lithography," *Proc. Kodak Microelectronic Seminar, Interface 82*, p. 139, Oct. 21–22, 1982.
[7] L. K. White and D. Meyerhofer, "Positive-Resist Processing Consideration for VLSI Lithography," *RCA Review*, **44**(1), p. 110 (1983).
[8] K. Bartlett, G. Hillis, M. Chen, R. Trutna, and M. Watts, "A Two-Layer Photoresist Process in a Production Environment," *SPIE Proc.*, **394**, p. 49 (1983).
[9] H. E. Abraham, K. G. Bartlett, G. L. Hillis, M. Stolz, and M. S. Wilson, *Hewlett-Packard Journal*, **34**(8), p. 34 (1983).
[10] M. Chen, W. R. Trutna, M. Watts, K. Bartlett, and G. Hillis, "Multilayer Photoresist Process Utilizing an Absorbant Dye," U.S. Patent 4,362,809, Dec. 7, 1982.
[11] M. Kaplan, D. Meyerhofer, and L. White, "Multilayer Resist Systems for VLSI Lithography," *RCA Review*, **44**(1), p. 135 (1983).
[12] M.Born and E. Wolf, *Principles of Optics*, MacMillan, New York (1964), 2nd ed., Chpt. 8.
[13] L. K. White, "Positive and Negative Tone Near-Contact Printing of Contact Hole Maskings," *RCA Review*, **43**(2), p. 391 (1982).
[14] W. G. Oldham, S. N. Nandgaonkar, A. R. Neureuther, and M. O'Toole, "A General Simulator for VLSI Lithography and Etching Processes: Part I—Application to Projection Lithography," *IEEE Trans. Electron Dev.*, **Ed-26**, p. 717 (1979).

An Analytical Study of the Winding Harmonics of a Saddle Deflection Coil

Basab B. Dasgupta
RCA Consumer Electronics, 600 N. Sherman Drive,
Indianapolis, IN 46201

Abstract—A closed-form analytical expression is derived for the harmonic (Fourier) coefficients of the winding distribution of a saddle-shaped magnetic-deflection coil by assuming that this distribution is expanded as a Fourier series at a given cross-section in terms of the polar angle about the coil axis. The dependence of the harmonics on various design parameters is analyzed. The effect of nonuniformity in the packing of wires during winding of the coil is discussed.

1. Introduction

The key to understanding the operating principles of wide-angle magnetic-deflection yokes used in modern color TV tubes is the harmonic analysis. In this analysis the current distributions in the horizontal and vertical coils are expanded as Fourier series in the polar angle θ about the yoke axis. Knowing the nth Fourier coefficients of the currents, one can determine the nth coefficients of the magnetic field components and hence the total magnetic field.[1,2] The harmonic coefficients of the wire distribution of a saddle coil, which is still in the design stage, are usually calculated by means of a computer program that Fourier analyzes the cross-section of the arbor cavity designed to make the proposed coil. While such programs are not difficult to run or understand, they do not provide physical insight into the dependence of the harmonic coefficients on the various geometric parameters of the cavity cross-section.

The purpose of the present paper is to derive an analytic expression for the harmonics of the winding distribution of *all* orders for a cavity-wound saddle coil. This expression allows us to make a

critical evaluation of the merits of any proposed design change in an existing arbor by analyzing its effect on various harmonics. It is also useful in designing a completely new arbor that would have a desired set of harmonics. Such an arbor design process is usually done by a computer program that goes through hundreds, possibly thousands, of iterations of changes in the design and calculates the harmonics in an attempt to match a meaningful design with the given harmonics. Obviously, if the computation of the harmonic coefficients is done analytically, the over-all program will be faster and more economical (if computer cost is a concern).

2. Calculation

A cross-section of a typical arbor cavity in a plane perpendicular to the coil axis is shown in Fig. 1, which also defines the various parameters involved (R, R_0, R_1, H, K, A, D, P, B and W). The saddle coil is made by winding turns through such a cavity and has a shape similar to the one shown in Fig. 2. Our derivation is based on the observation that typically the H and K parameters of the cavity are

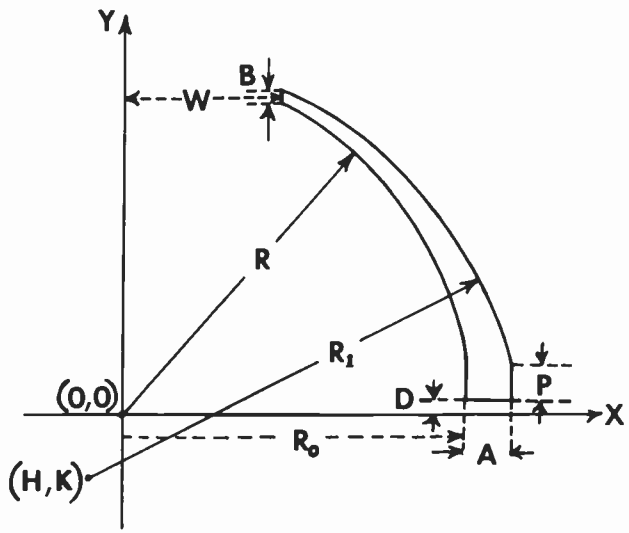


Fig. 1—Cross-section of the arbor cavity in a plane perpendicular to the coil axis; only one quadrant is shown. Its boundary consists of two circular arcs, one of radius R and center at the origin (0, 0) and the second one of radius R_1 and center at the point (H, K), plus linear segments of lengths B, A, and P.

small* compared to R_1, and hence one can make binomial expansions of the relevant expressions in powers of the ratios H/R_1 and K/R_1. For the sake of simplicity, we assume that the arbor is packed uniformly by wires during the winding process so that the harmonics of the wire distribution are the same as those of the incremental area variation of the arbor cross-section (apart from a constant multiplicative factor).

We will divide the cavity cross-section into five regions, as shown in Fig. 3, by different labelings. If $f(\theta)d\theta$ is the incremental area of the cavity between the angles θ and $\theta + d\theta$, then it can be easily shown[3] that $f(\theta)$ is given by the expression

$$2f(\theta) = \begin{cases} (R_0 + A)^2 \sec^2\theta - D^2\mathrm{cosec}^2\theta & \text{in Region 1} \\ (R_0 + A)^2 \sec^2\theta - R_0^2\sec^2\theta & \text{in Region 2} \\ g(\theta) - R_0^2\sec^2\theta & \text{in Region 3} \\ g(\theta) - R^2 & \text{in Region 4} \\ g(\theta) - W^2\sec^2\theta & \text{in Region 5} \end{cases} \quad [1]$$

where $g(\theta)$ is given by

$$g(\theta) = (H^2 - K^2)\cos 2\theta + 2HK\sin 2\theta + R_1^2 + 2(H\cos\theta + K\sin\theta)\{R_1^2 - (H\sin\theta - K\cos\theta)^2\}^{\frac{1}{2}}. \quad [2]$$

Using the binomial theorem, we can write

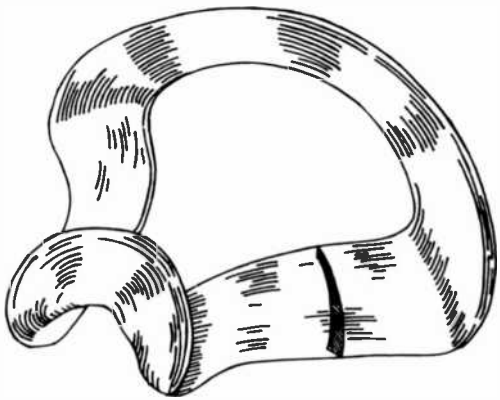


Fig. 2—One-half of a saddle coil wound in an arbor cavity. A cross-section is shown as the shaded area; this should be identical to Fig. 1 in each plane perpendicular to the coil axis if wires are uniformly packed into the cavity.

* For almost all commercial coils, R_1 is typically in the range of one to three inches, depending on the tube geometry and coil cross section, whereas H and K are of the order of 0.2 inch or less.

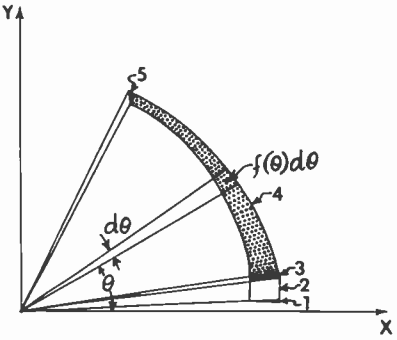


Fig. 3—The cavity cross-section can be divided into five regions depending upon the shape.

$$\{R_1^2 - (H\sin\theta - K\cos\theta)^2\}^{\frac{1}{2}} \cong R_1 - (H\sin\theta - K\cos\theta)^2/2R_1, \quad [3]$$

so that

$$g(\theta) = R_1^2 + (H^2 - K^2)\cos 2\theta + 2HK\sin 2\theta + 2R_1(H\cos\theta + K\sin\theta) - (H\cos\theta + K\sin\theta)(H\sin\theta - K\cos\theta)^2/R_1. \quad [4]$$

The nth Fourier component of $f(\theta)$ is given by*

$$f_n = (4/\pi)\int_0^{\pi/2} f(\theta)\cos n\theta d\theta. \quad [5]$$

It is now a straight-forward but somewhat tedious exercise to substitute Eqs. [1] and [4] into Eq. [5] and carry out the indicated integration. The resulting expression for f_n is inconveniently long. To write it in a compact form let us introduce the following quantities

$$S_{nm}(\theta) \equiv \frac{\sin(n+m)\theta}{2(n+m)} + \frac{\sin(n-m)\theta}{2(n-m)} \quad [6]$$

$$C_{nm}(\theta) \equiv \frac{\cos(n-m)\theta}{2(n-m)} - \frac{\cos(n+m)\theta}{2(n+m)} \quad [7]$$

$$I_n(\theta) \equiv \int \frac{\cos n\theta}{\cos^2\theta} d\theta$$

$$= (-1)^{(n-1)/2} n \ln\tan\left(\frac{\pi}{4} + \frac{\theta}{2}\right)$$

* In Eq. [5] as well as in Eq. [20], we impose the sign convention that $f(\theta)$ (and $A(\theta)$) are positive in the first and fourth quadrants (where the current in the coil is going in one direction) and negative in the other two quadrants (where the current goes in the opposite direction).

$$+ 4(1 - \delta_{n,1}) \sum_{k'=0}^{(n-3)/2} \binom{n-2}{k'} \frac{\sin(n - 2k' - 2)\theta}{(n - 2k' - 2)}$$

$$+ 4(1 - \delta_{n,1})(1 - \delta_{n,3}) \sum_{k=1}^{(n-3)/2} \sum_{k'=0}^{(n-3-2k)/2} (-1)^k \frac{n}{k}$$

$$\binom{n-k-2}{k'} \binom{n-k+1}{k-1} \frac{\sin(n - 2k - 2k' - 2)\theta}{(n - 2k - 2k' - 2)}, \quad [8]$$

and

$$J_n(\theta) \equiv \int \frac{\cos n\theta}{\sin^2\theta} d\theta$$

$$= 2^{(n-1)} \sum_{k=0}^{(n-1)/2} (-1)^k \binom{(n-1)/2}{k} \frac{\sin^{2k-1}\theta}{(2k - 1)}$$

$$+ (1 - \delta_{n,1}) \sum_{k=1}^{(n-1)/2} \sum_{k'=0}^{(n-1-2k)/2} (-1)^{k+k'} \frac{n}{k} 2^{n-1-2k}$$

$$\binom{(n-1-2k)/2}{k'} \binom{n-k+1}{k-1} \frac{\sin^{2k'-1}\theta}{(2k' - 1)}. \quad [9]$$

In Eqs. [8] and [9] $\delta_{n,m}$ is defined as

$$\delta_{n,m} = \begin{cases} 1 \text{ if } n = m \\ 0 \text{ if } n \neq m \end{cases} \quad [10]$$

and

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}. \quad [11]$$

We can now write the expression for f_n as

$$f_n = (2/\pi)[(R_0 + A)^2\{I_n(\theta_3) - I_n(\theta_1)\} - D^2\{J_n(\theta_2) - J_n(\theta_1)\}$$
$$- R_0^2\{I_n(\theta_4) - I_n(\theta_2)\} + R_1^2\{S_{n0}(\theta_6) - S_{n0}(\theta_3)\}$$
$$+ (H^2 - K^2)\{S_{n2}(\theta_6) - S_{n2}(\theta_3)\}$$
$$+ 2HK\{C_{n2}(\theta_6) - C_{n2}(\theta_3)\} + \{2R_1H - H(H^2 + K^2)/8R_1\}$$
$$\{S_{n1}(\theta_6) - S_{n1}(\theta_3)\}$$
$$+ \{2R_1K - K(H^2 + K^2)/8R_1\}\{C_{n1}(\theta_6) - C_{n1}(\theta_3)\}$$
$$- \frac{H(3K^2 - H^2)}{4R_1}\{S_{n3}(\theta_6) - S_{n3}(\theta_3)\}$$

$$- \frac{K(K^2 - 3H^2)}{4R_1}\{C_{n3}(\theta_6) - C_{n3}(\theta_3)\}$$

$$- R^2\{S_{n0}(\theta_5) - S_{n0}(\theta_4)\} - W^2\{I_n(\theta_6) - I_n(\theta_5)\}], \quad [12]$$

where

$\theta_1 = \arctan\{D/(R_0 + A)\}$
$\theta_2 = \arctan(D/R_0)$
$\theta_3 = \arctan\{(D + P)/(R_0 + A)\}$
$\theta_4 = \arctan\{(D + P)/R_0\}$
$\theta_5 = \arccos(W/R)$
$\theta_6 = \arctan\{(R\sin\theta_5 + B)/W\}$

The expression given by Eq. [12] is accurate up to terms of order $(H/R_1)^3$ and $(K/R_1)^3$, but the binomial expansion used in Eq. [2] can be continued to include higher order terms and an expression can be derived that is accurate to any desired degree of accuracy.

3. Dependence on Design Parameters

This section will be devoted to a discussion of the effect of varying certain geometric cavity parameters on the various harmonics.

3.1 Dependence on *H*, *K*, and *W*

In order to analyze this dependence let us assume that $D = P = 0$ and $\theta_6 \cong \theta_5 = \theta_0$ so that $W \cong R\cos\theta_0$; this assumption simplifies the analysis and should not significantly change our conclusions. Furthermore, let us retain only terms up to second power in H and K. Then Eq. [12] reduces to

$$f_n = (2/\pi)[(R_1^2 - R^2)(\sin n\theta_0)/n + (H^2 - K^2)S_{n2}(\theta_0) + HK\{C_{n2}(\theta_0) - C_{n2}(0)\} + 2R_1HS_{n1}(\theta_0) - 2R_1K\{C_{n1}(\theta_0) - C_{n1}(0)\}] \quad [13]$$

Several special cases of this equation deserve attention.

Case (1): $H = K = 0$
This case corresponds to an arbor cavity consisting of two concentric circles. For such a cavity, any harmonic can be made to vanish by an appropriate choice of θ_0. For example, f_n will be zero if θ_0 is chosen to be π/n.

Case (2): $K = 0, R = R_1, \theta_0 \cong \pi/2$ and H is small (so that its square can be neglected

Then Eq. [13] reduces to

$$f_n = \begin{cases} 0 & \text{if } n \neq 1 \\ R_1 H & \text{if } n = 1 \end{cases} \qquad [14]$$

i.e., the cavity would produce a coil with an almost pure cosine winding distribution (note that θ_0 cannot be exactly $\pi/2$ because some nonzero window length is required for winding the coil). This "trick" has been used in the past by yoke designers to design saddle coils that produce uniform fields.

Case (3): $n = 3$ and $40° < \theta_0 < 70°$
This case is interesting because in a self-converging saddle-toroidal yoke, the angle θ_0 in the saddle coil is typically of the order of 40° near the front end of the yoke and 70° near the rear end and varies monotonically in between. Furthermore, the value of the third harmonic coefficient f_3 is the most significant in determining the yoke performance. From Eq. [13], we see that

$$f_3 \cong (2/\pi)[(R_1^2 - R^2)\sin 3\theta_0/3 + R_1 H(\sin 4\theta_0 + 2\sin 2\theta_0)/4 - R_1 K(2\cos 2\theta_0 - \cos 4\theta_0 - 1)/4]. \qquad [15]$$

If one plots the coefficients of H and K as functions of θ_0, it is easy to see that the coefficient of H becomes large and that of K becomes small as θ_0 approaches 40°; the converse is true as θ_0 approaches 70°. The $\sin 3\theta_0$ term is zero at $\theta_0 = 60°$ and, in general, small near the rear-end; also the converging properties of a yoke are *relatively* insensitive to the winding distributions near the front-end of the saddle coil (where the effect of H is largest). We conclude that *K is more critical than H as a variable design parameter in making a satisfactory yoke*. This observation is consistent with the traditional design practice of "opening" or "closing" the cavity, which is effectively a way of varying the K parameter.

Case (4):
The change in the harmonics caused by the addition of one turn of angular diameter $\Delta\theta (= d/R$, d being the actual wire diameter) at the "window" (i.e., at angle θ_0) can be expressed as

$$\Delta f_n \cong (\partial f_n/\partial\theta_0)\Delta\theta = (2\cos n\theta_0 \Delta\theta/\pi)[(R_1^2 - R^2) + (H^2 - K^2)\cos 2\theta_0 + HK\sin 2\theta_0 + 2R_1 H\cos\theta_0 + 2R_1 K\sin\theta_0] \qquad [16]$$

The change in the n-th harmonic is most pronounced if $\theta_0 = \pi/n$ and zero if $\theta_0 = \pi/2n$

Case (5): n is very large
If $n \gg 2$, the leading term in Eq. [13] takes the form

$$f_n = (2/\pi)[(R_1^2 - R^2) + (H^2 - K^2) + R_1 H] \sin n\theta_0/n$$
$$\approx (2/\pi)[(R_1 + H/2)^2 - R^2] \sin n\theta_0/n \qquad [17]$$

This equation tells us that *for very high order harmonics the cavity behaves as if it consisted of two concentric circles of radii R and $R_1 + H/2$.*

3.2 Dependence on D, P, and A

Like H and K, these parameters are typically small (compared to R or R_0). If we subtract the expression given by Eq. [13] (which is valid when $D = P = 0$) from the one given by Eq. [12] and retain only terms of the lowest order in the small quantities D, P, and A, we get the following correction term:

$$\Delta f_n \cong (2/\pi)[-AD + AP - (R_1^2 - R^2)(D + P)/R_0]. \qquad [18]$$

The most significant thing about this equation is that Δf_n is independent of n and each parameter is involved linearly. Hence we can draw the following conclusion which might be a useful guideline to yoke designers:

If any design parameter near the x-axis (D, P, or A) is changed by a small amount while the others are kept constant, it changes all harmonics by the same amount (i.e. smaller harmonics undergo larger % changes), and the "slope" of this change (i.e. change in harmonic per unit change in the value of the parameter) is the same for all harmonics.

4. Effect of Nonuniform Packing of Wires

The discussion up to this point has been based on the assumption that the arbor cavity is packed uniformly by wires during the winding process. This is seldom the case in reality because of mechanical variations in the winding machine. If such variations are completely random then it is likely that the effect of nonuniformity in the wire packing would be averaged out over the length of the coil. The purpose of this section is to analyze the effect of systematic variations in the packing on the harmonics.

First, let us define a "packing factor function" $p(\theta)$ as follows: let

$A(\theta)d\theta$ be the area of the cavity cross-section between the angles θ and $\theta + d\theta$ and let $f(\theta)d\theta$ be the area actually occupied by the wire cross-sections; then define $p(\theta)$ such that

$$f(\theta) = p(\theta)\, A(\theta). \qquad [19]$$

If $p(\theta)$ arises because of systematic variations in the winding machine, then it would be symmetric about *both* x and y axes. Furthermore, because of our sign convention*, $p(\theta)$ must always be positive. Hence it can be expanded as†

$$p(\theta) = \sum_{n=0,2,4\ldots} p_n \cos n\theta. \qquad [20]$$

$A(\theta)$ can be expanded as usual:

$$A(\theta) = \sum_{n=1,3,5\ldots} A_n \cos n\theta. \qquad [21]$$

Substituting Eqs. [19], [20], and [21] into Eq. [5], we get

$$f_n = p_0 A_n + \sum_{m \neq n} (p_{m+n} + p_{|m+n|}) A_m/2. \qquad [22]$$

The first term, of course, corresponds to a uniform packing and $|p_0| \gg |p_n| (n > 0)$. An important general result following from this equation can be stated as a theorem: *If $p(\theta)$ has a large value for its harmonic coefficient p_l, where l is some specific even integer, then f_n and A_n can be considerably different when $n = l \pm 1$.*

We will discuss a simple but important special case in more detail. Consider the packing represented by the function

$$p(\theta) = p_0 + p_2 \cos 2\theta, \qquad [23]$$

with $|p_0| \gg |p_2|$. It corresponds to a gradual decrease (if $p_2 > 0$) or increase (if $p_2 < 0$) of the packing factor with angle θ from the value $(p_0 + p_2)$ at $\theta = 0°$ to $(p_0 - p_2)$ at $\theta = 90°$. The first few winding harmonics in this case are given by

$$f_1 = p_0 A_1 + p_2 A_3/2 \qquad [24a]$$

$$f_3 = p_0 A_3 + p_2 (A_1 + A_5)/2 \qquad [24b]$$

* See footnote to Eq. [5].

† Here, we assume that although the packing is nonuniform, it is still symmetric about $x - z$ and $y - z$ planes. One can further generalize our treatment by taking into account various types of unsymmetries in wire packing.

$$f_5 = p_0 A_5 + p_2(A_3 + A_7)/2 \qquad [24c]$$

$$f_7 = p_0 A_7 + p_2(A_5 + A_9)/2 \qquad [24d]$$

We note that $f_n \approx p_0 A_n$ for all n except $n = 3$. For $n = 3$, the $p_2 A_1$ term in Eq. [24b] can be comparable to $p_0 A_3$ even though $|p_2| \ll |p_0|$ if $|A_3| \ll |A_1|$. This latter condition may indeed be valid near the rear end of most yokes; since the converging properties of a yoke are rather sensitive to the value of f_3 near the rear-end, a nonzero value of p_2 must be taken into account in computing the yoke characteristics. As we approach the front end of the coil, A_3/A_1 increases and the term $p_0 A_3$ dominates in Eq. [24b] and the variable nature of $p(\theta)$ does not make too much difference in yoke performance. A similar conclusion can be drawn regarding f_3 if $p(\theta)$ has a nonzero p_4.

If $p(\theta)$ has higher order coefficients, such as $p_6, p_8, \ldots$, then the wire harmonics f_n and cavity harmonics A_n may be quite different near the rear end of the coil (for $n = 5, 7 \ldots$). But this observation is inconsequential, since the electron beam near the rear lies very close to the axis where the higher harmonics of the magnetic field are negligible anyway.

In a practical situation, p_0, p_2, and p_4 may be determined for a given type of coil and a given winding machine by digitizing a number of cross-sections of a coil made by this machine and comparing the wire harmonics determined by digitization with the cavity harmonics. One can then hope that for any other coil of this "type", the coefficients p_0, p_2, p_4 would remain approximately the same.

5. Conclusion

We have derived an analytical expression for the winding harmonics in a saddle coil, which can be either the horizontal-deflecting or vertical-deflecting coil of a magnetic-deflection yoke. Several general conclusions have been drawn regarding the relationship between the harmonics and the design parameters of the coil on the basis of this expression. Some of these conclusions justify the traditional manipulations of the yoke designers, while the others might be useful in understanding yet unsolved mysteries about yokes.

References:

[1] D. M. Fye, "An Integral Equation Method for the Analysis of Magnetic Deflection Yokes," *J. Appl. Phys.*, **50**, p. 17, Jan 1979.
[2] Y. Yokota and T. Toyofuku, "The Calculation of the Deflection Magnetic Field and the Electron-Beam Trajectory for Color Television," *IEEE Trans. Consumer Electronics*, **25**, p. 91 (1979).
[3] S. V. Kissinger, private communication.

Extending the Content and Expanding the Usefulness of the Simple Gaussian Lens Equations—Part 3, Section B: (c) (Continued) Permissible Sign Groupings That Typify Non-Afocal Relaying Actions Generally; (d) The Single Thick Lens Treated as a System of Two Quasi-Separated Lenses; (e) Concluding Summary Statement.

L. T. Sachtleben
RCA Advanced Technology Laboratories, Camden, NJ

Abstract—The Gaussian expressions for the optical separation of the two lenses of a non-afocal primitive relay optical system are developed in general forms that are suited to any one of the four conjugate distances being assigned an infinite value. These are extended to the case where the members of one of the conjugate pairs are both infinite, and the system afocal. The afocal system is analyzed to obtain relationships that simplify Gaussian design of an afocal telescope when *all four* conjugate distances are finite. Examples are worked through to illustrate applications. The relaying action of *any* non-afocal optical system (for two discrete object distances) is studied in its most general terms, with the four conjugate distances, the focal length, and the two magnifications being represented as literal parameters with fixed signs. Eight optically possible sign groupings of the parameters are identified when conjugate distances are Newtonian. They are used to identify the eighteen optically possible sign groupings when the conjugate distances are Gaussian. The two sorts of groupings are tabulated for easy reference and to simplify their intercomparison. Analyses that relate the Newtonian and Gaussian groupings are performed in terms of inequalities. Examples of possible applications are reviewed. Useful applications of the theory of two separated lenses to the problems of a single thick lens are developed. Examples are worked through.

29. Continuation of Part 3, Section A, *RCA Review* 45, No. 1, pp. 136–149.

V. The Eighteen Optically Possible Sign-Groupings of the Ten Gaussian Parameters; In *Any* Centered, Non-Afocal System of Spherical Lenses, In Air or Vacuum, When Used As a Relay†

The construction of Table 15 has been the objective toward which the studies of this section (Sec. 29) have been consistently directed. In particular, Appendix 9 describes the final procedure for constructing Table 15 from Table 14; it may be studied separately.

Table 15 provides a basis for preliminarily judging the problem of designing a non-afocal relaying optical system (for use in air or vacuum) when the problem is in its earliest, formative, or proposal stage. It is assumed that at this stage, the proposed or prospective system is defined *solely* in terms of the signs of some or all ten of the Gaussian parameters of columns (1), (2)*, (3)′, . . . , (8), (10), and (11) of the table.

For example, if the proposed signs of the ten parameters are all given, or assumed, this set of signs *must* correspond *exactly* to that of one of the eighteen groups of Table 15. Otherwise, at least some of the signs must be revised until such a correspondence exists for in the absence of such a correspondence the proposed relaying system will be impossible to design.

If the signs of *fewer* than the ten Gaussian parameters are given, or assumed, and if this partial set of signs corresponds *exactly* to that of the corresponding parameters of one or more of the eighteen groups of Table 15, then the signs of the remaining parameters must be made the same as those of one or another of such groups. Otherwise the prospective system is not designable, unless revised.

Table 15 applies just as well to the case of a system that produces a single image of a single object, as to that of a system that images two axially separated objects. For the case of a single object Table 15 simplifies, being then restricted to Column Nos. (1), (2)*, (4)*, and (6)* and to Group Nos. I_a, I_d, IV, V, $VIII_a$, and $VIII_d$. Columns

† Editors Note: This paper is a continuation and conclusion of the author's previous papers in the December 1976, June 1978, and March 1984 issues of the *RCA Review*. Section and equation numbers continue from Part 3, Section A.

Table 15—Schedules of the eighteen possible sign groupings† of the ten Gaussian parameters of columns (1) through (8), (10), and (11) in *any* centered, non-afocal system of spherical lenses in air or vacuum when used as a relay (see Appendix 9). X, X', Y, and Y' (without subscripts), *always* identify Gaussian conjugate distances that are measured from the system's two principal points. System numbers correspond to those of Table 14 (see Appendix 9).

Column Nos.:		(1)	(2)*	(3)′	(4)*	(5)′	(6)*	(7)′	(8)	(9)	(10)	(11)
Parameters		$f \gtrless 0$ Finite	X	X'	Y	Y'	M	M'	$M_\Delta = MM' = \Delta_c y/\Delta_c x$	Relaying Action	$\Delta_c x$ Or	$\Delta_c y$ Or
In Reference To:		Focal Length	Object O	Object O′	Image I	Image I′	Mag. O to I	Mag. O′ to I′	Displacement Mag.	Regular Or Inv.	Displt. $(X' - X)$	Displt. $(Y' - Y)$
Group & Sys. Nos.												
I_a.	1 & 1′	P	N	N	N	N	P > 1	P > 1	P	Regular	P/N	P/N
I_b.	1 & 2′	P	N	P	N	P	P > 1	P < 1	P	Regular	P	P
I_c.	2 & 1′	P	P	N	P	N	P < 1	P > 1	P	Regular	N	N
I_d.	2 & 2′	P	P	P	P	P	P < 1	P < 1	P	Regular	P/N	P/N
II_a.	1 & 5′	P	N	N	N	P	P > 1	N	N	Inverse	N	P
II_b.	2 & 5′	P	P	N	P	P	P < 1	N	N	Inverse	N	P
III_a.	5 & 1′	P	N	N	P	N	N	P > 1	N	Inverse	P	N
III_b.	5 & 2′	P	N	P	P	P	N	P < 1	N	Inverse	P	N
IV.	5 & 5′	P	N	N	P	P	N	N	P	Regular	P/N	P/N

Table 15—Continued

Column Nos.:		(1)	(2)*	(3)′	(4)*	(5)′	(6)*	(7)′	(8)	(9)	(10)	(11)
Parameters In Reference To:		$f \gtrless 0$ Finite Focal Length	X Object O	X' Object O′	Y Image I	Y' Image I′	M Mag. O to I	M' Mag. O′ to I′	$M_\Delta = MM' = \Delta_c y/\Delta_c x$ Displacement Mag.	Relaying Action Regular Or Inv.	$\Delta_c x$ Or Displt. $(X' - X)$	$\Delta_c y$ Or Displt. $(Y' - Y)$
V.	6 & 6′	N	P	P	N	N	N	N	P	Regular	P/N	P/N
VI_a.	6 & 4′	N	P	N	N	N	N	P < 1	N	Inverse	N	P
VI_b.	6 & 3′	N	P	P	N	P	N	P > 1	N	Inverse	N	P
VII_a.	4 & 6′	N	N	P	N	N	P < 1	N	N	Inverse	P	N
VII_b.	3 & 6′	N	P	P	P	N	P > 1	N	N	Inverse	P	N
$VIII_a$.	4 & 4′	N	N	N	N	N	P < 1	P < 1	P	Regular	P/N	P/N
$VIII_b$.	4 & 3′	N	N	P	N	P	P < 1	P > 1	P	Regular	P	P
$VIII_c$.	3 & 4′	N	P	N	P	N	P > 1	P < 1	P	Regular	N	N
$VIII_d$.	3 & 3′	N	P	P	P	P	P > 1	P > 1	P	Regular	P/N	P/N

† See final paragraph of Appendix 9.

(3)′, (5)′, and (7)′ are superfluous, and Columns (8), (9), (10), and (11) do not apply in the case of a single object; also, all the primed system numbers in the left hand column of Table 15 are superfluous.†

By way of review of the eighteen groups, I_a through $VIII_d$ of Table 15, all share the following important internal relationships, which *must* hold true within each of them. Parameters X, X', Y, and Y' are measured from the two principal points of the system to the four object- and image-points; the locations of all six of these points may be partly or wholly unknown. Nevertheless, the following relationships *must* effectively apply in the case of any group of Table 15.

$M = \dfrac{Y}{X}$ and $M' = \dfrac{Y'}{X'}$; see Table I (V-1), Part 1.

$\Delta_c x = X' - X$ and $\Delta_c y = Y' - Y$; see Sec. 10, Part 1, Sec. 17, Part 2, and Eqs. [172] and [173], Part 2.

$$\frac{XY}{X - Y} = \frac{X'Y'}{X' - Y'} = f \text{ (or } F\text{)}; \text{ see Table 1 (II-1), Part 1.}$$

$$\frac{-MM'\Delta_c x}{M' - M} = \frac{-\Delta_c y}{M' - M} = f \text{ (or } F\text{)}; \text{ see Eqs. [42] and [43], Part 1.}$$

$\Delta_c y/\Delta_c x = MM' = M_\Delta$; see Eq. [41], Part 1, and Eq. [232], Part 3.

Any of the foregoing relationships may be expressed in many other useful ways, as suggested by the 33 given or implied relationships that are grouped in Table 1, Part 1.

When the system of any group, I_a through $VIII_d$ of Table 15, has the form of a separated two-lens system, it may be necessary to find the values of f_1, f_2, and their optical separation a. Sometimes one or another of these three values may be given as a problem requirement, or it may be assigned an arbitrary trial value or a set of such values. In solving for these three quantities the solutions may take numerous forms depending on the nature of the problem. The solution may contain any of the parameters of Table 15, and in addition may include any of the parameters D, D', D_M, D'_M, x_1, x'_1, y_2,

† When considering a single object and its image at conjugate distances X and Y, the usefulness of the entire Table 15 may be preserved by the fictional device of assuming X' and Y' each to be appropriately related to and arbitrarily close to zero, while also assuming M' to be appropriately related to and arbitrarily close to $+1$. The reference to Table 14 then also continues to apply to the primed system numbers of Table 15; while $\Delta_c x = X' - X = -X$, $\Delta_c y = Y' - Y = -Y$, and $M_\Delta = MM' = M$.

y_2', S_1, S_2, and $\overline{S}$ of Fig. 12. Numerous expressions throughout the three parts of this paper involve some of these quantities; they may be important to formulating the solution, as will be suggested by the following examples. The variety of forms that problems may take and the number of solutions that may be required are great and a few examples cannot develop them all (see last paragraph of Sec. 7, Part 1, and, in particular, all of Part 2).

29A. Representative Applications

Example 13. A certain image-relaying problem can apparently be solved by making $X > 0$ and $Y < 0$, and at the same time making $X' < 0$ and $Y' > 0$. The advisability of attempting a solution along this line calls for Table 15 to be consulted. Should development of such a system be attempted?

The schedule of Table 15 shows this to be an impossible condition to meet, because the signs of X and Y call for the system focal length to be negative whereas the signs of X' and Y' require it to be positive. The only system that could meet these conflicting requirements would have to have a "zero" focal length, which is impossible (see closing discussion in Appendix 7),† or it would have to have an infinite focal length. The latter possibility should be studied further in the light of Secs. 26 and 26A. Table 15 does not apply to afocal systems, and the matter is not considered further here.

Example 14. This example does not attempt to represent any particular practical problem or application, but rather to indicate how the parameters may be related among themselves in any group of Table 15, including parameters not explicitly appearing in that table. It also leads to consideration of how a simple change in the value of a parameter in one group may change a problem to membership in another group.

The signed numerical values of three of the parameters of Table 15 are given the values shown in Table 16, where they are arranged (left to right) according to the general scheme of Table 15. Quantities not known are indicated by "?". The signs of the given quantities appear in line 4 of the table. The units chosen for the linear quantities may be arbitrary.

From Table 16, all of the remaining seven unknown quantities are readily calculated as follows. (Note: The final signs of all quantities in the table, as determined below, are shown in line 5; they

† Consult final paragraphs of Appendix 7 and Fig. 21.

identify the system as corresponding to Group IV, Table 15. The signs in the last two columns need only be alike; either both P or both N.)

By (V-1), Table 1, $M = \dfrac{Y}{X} = \dfrac{84}{-12} = -7$. Col. (6)*.

By (II-1), Table 1, $F = \dfrac{XY}{X - Y} = \dfrac{-12(84)}{-12 - 84} = +10.500000.$ Col. (1).

By (III-5), Table 1, $X' = \dfrac{FY'}{F - Y'} = \dfrac{10.5(17)}{10.5 - 17} = -27.461538.$ Col. (3)'.

By (V-1), Table 1, $M' = \dfrac{Y'}{X'} = \dfrac{17}{-27.461538} = -0.619048.$ Col. (7)'.

By Eq. [232], Sec. 27, $M_\Delta = MM' = -7(-0.619048)$
$= +4.333333.$ Col. (8).

By Eq. [172], Part 2, $\Delta_c x = X' - X = -27.461538 - (-12)$
$= -15.461538.$ Col. (10).

By Eq. [173], Part 2, $\Delta_c y = Y' - Y = 17 - 84$
$= -67.000000.$ Col. (11).

The positive value of M_Δ indicates the relaying action is "regular," as recorded in Col. (9), line 5.

The signs of the calculated values of the seven unknown (or "?") parameters of Table 16 match Group IV, Table 15. The signs in columns (10) and (11) may be either P or N, but must both be the same; in this case, both negative. If the signs in line 4 had been *all* that was known of the system, then the system could have corresponded to either Group IIIb or Group IV, Table 15, but to no others.†

Table 15 also suggests that by suitably changing the sign and numerical value of Y', the signs of all the parameters could be made to match Group III_a of the table; and that by similarly changing X' the signs could be made to match Group III_b. These suggested possibilities will be more fully discussed in Example 15.

Figure 12 (repeated from Part 3, Sec. A for the reader's convenience) schematically represents a non-afocal, two-lens relay system and identifies all of the parameters of Table 15. Additional

† Readers should examine the overall effect on Table 16, including the effect on the group-number of its Table 15 system, of changing the Y- quantity to $Y = 8$; also of changing the Y-quantity to $Y = -8$. No changes are to be made in the given values of X and Y'.

Table 16—Three typical parameters from Table 15 as a basis for the study of Example 14

Line No.	Subject											
1	Column No.	(1)	(2)*	(3)′	(4)*	(5)′	(6)*	(7)′	(8)	(9)	(10)	(11)
2	Parameter	F	X	X'	Y	Y'	M	M'	M_Δ	Relaying Action	$\Delta_c x$	$\Delta_c y$
3	Value	?	−12	?	84	17	?	?	?	?	?	?
4	Sign	?	N	?	P	P	?	?	?	?	?	?
5	Final Sign	P	N	N	P	P	N	N	P	Reg.	N	N

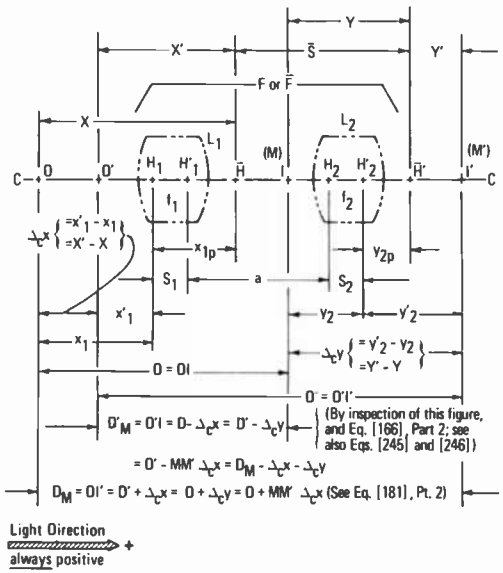


Fig. 12—Quantities involved in the Gaussian design relationships for the primitive relay optical system. Similar to Fig. 10, Part 2. Schematic only—not to scale.

parameters included in the figure are the D, D', and D_M of Fig. 10, Part 2; they are usefully related to the other parameters by Eqs. [181], [182], [183], etc., of Part 2. A new parameter D_M' is included in the figure. Its value is indicated by inspection of the figure to be

$$D_M' = D' - \Delta_c y = D - \Delta_c x = D' - MM'\Delta_c x \qquad [245]$$

from which, with the aid of Eq. [181], we may write

$$D_M' = D_M - (\Delta_c x + \Delta_c y). \qquad [246]$$

In Fig. 12, the principal points of the system are identified as $\overline{H}$ and $\overline{H}'$, respectively, at distances x_{1p} and y_{2p} from the first and second principal points of lenses L_1 and L_2. They are separated by the distance $\overline{S}$. The location of the object- and image-points with reference to the pair of separated lenses is indicated by the Gaussian object-distances X and X', measured from $\overline{H}$, and by the corresponding image-distances Y and Y', measured from $\overline{H}'$. The object- and image-locations are also indicated, respectively, by distances x_1 and x_1', measured from the first principal point of lens L_1, and by y_2 and y_2', measured from the second principal point of lens L_2. The original numerical subscripts of Part 1, Secs. 5–9, are used, but

they are subject to omission according to the scheme of Part 2, Sec. 18, if the reader prefers.

By inspection of Fig. 12, we write the general equation for a non-afocal, separated two-lens system in terms of the parameters for object O and its image I. This equation is

$$D = -X - x_{1p} + S_1 + a + S_2 + y_{2p} + Y. \qquad [247]$$

Adding primes to D, X, Y, x_{1p}, and y_{2p} in this equation, makes it apply to object O′ and its image I′. The relationships of continued Eq. [181] permit Eq. [247] to be easily expressed in terms of D_M and $\Delta_c y$, or of D_M and $\Delta_c x$, should this be desired.

By multiplying the right side of Eq. [32], Part 1, by $f_2/f_2 = 1$, a substitution from Eq. [45] (or [113]) permits us to write

$$x_{1p} = \frac{aF}{f_2}. \qquad [248]$$

A similar procedure applied to Eq. [33], where the multiplying factor is $f_1/f_1 = 1$, yields the equation

$$y_{2p} = \frac{-aF}{f_1}. \qquad [249]$$

Equation [45] is symmetrical in f_1 and f_2; its solution for f_2 is

$$f_2 = \frac{(f_1 - a)F}{f_1 - F}, \qquad [250]$$

in which the just-mentioned symmetry property allows the numerical subscripts to be interchanged, if desired, to obtain the solution of Eq. [45] for f_1.

When the expressions for x_{1p}, y_{2p}, and f_2 of Eqs. [248], [249], and [250] are substituted into Eq. [247] to eliminate those three variables, the resulting equation simplifies to the standard form of a quadratic in f_1, which is

$$f_1^2(D - Y + X - S_1 - S_2) - f_1 a[(D - Y + X - S_1 - S_2) - a] - a^2F = 0, \qquad [251]$$

from which the solution for f_1 is

$$f_1 = \frac{a[(D - Y + X - S_1 - S_2) - a]}{2(D - Y + X - S_1 - S_2)} \pm \left[\frac{a^2[(D - Y + X - S_1 - S_2) - a]^2}{4(D - Y + X - S_1 - S_2)^2} + \frac{a^2F}{D - Y + X - S_1 - S_2}\right]^{1/2}. \qquad [252]$$

For solutions to be real, the value of the discriminant of Eq. [251] must be

$$a^2[(D - Y + X - S_1 - S_2) - a]^2 + 4a^2F(D - Y + X - S_1 - S_2) \geq 0. \quad [253]$$

The values of X, Y, and F in Eq. [252] have already been given or calculated. The values of $a > 0$, D, S_1, and S_2 may be assigned arbitrarily, except that the values of the seven quantities must together satisfy Ineq. [253]. The seven values, which are given in Table 16, are calculated immediately following it, or are arbitrarily assumed, are

$X = -12$ Given in Table 16
$Y = 84$ Given in Table 16
$F = 10.5$ Calculated following Table 16
$a = 5.0$ Arbitrary
$D = 48$ Arbitrary
$S_1 = 0.25$ Arbitrary
$S_2 = 0.25$ Arbitrary

Additional values to be used in this example are calculated immediately following Table 16 as

$M = -7$
$X' = -27.461538$
$M' = -0.619048$
$M_\Delta = MM' = +4.333333$
$\Delta_c x = -15.461538$
$\Delta_c y = -67.000000$

By substitutions into the left side of Ineq. [253] of the values of the first seven parameters that follow it,

$$5^2(48 - 84 - 12 - 0.5)^2 + 4(5)^2(10.5)(48 - 84 - 12 - 0.5) = 20631.25.$$

This satisfies the inequality and assures that the solution for f_1, by Eq. [252], will be real. Accordingly, by substitutions into Eq. [252]

$$f_1 = 2.757732 \pm 1.480782 = 4.238514 \text{ or } 1.276950,$$

and by substitutions into Eq. [250], the corresponding solutions for f_2 are

$$f_2 = 1.276950 \text{ or } 4.238514,$$

which, due to the symmetry of the equations in f_1 and f_2, serves as

a check on the values of f_1. As a further check, using the relationship of Eq. [45], Part 1,

$$F = \frac{f_1 f_2}{f_1 + f_2 - a} = \frac{4.238514(1.276950)}{4.238514 + 1.276950 - 5.0}$$
$$= 10.500000, \text{ a perfect check!}$$

Numerous further checks are of interest, one of them by Eq. [180], Part 2, when expressed in the form

$$a = \frac{M(x_1' - \Delta_c x)(y_2 + MM'\Delta_c x) - M'x_1'y_2}{MM'\Delta_c x}. \qquad [254]$$

Here, $MM'\Delta_c x = -67.000000$. Also, by Eq. [248] and inspection of Fig. 12,

$$x_1' = X' + \frac{Fa}{f_2}, \qquad [255]$$

and by Eq. [249] and Fig. 12,

$$y_2 = Y - \frac{Fa}{f_1}. \qquad [256]$$

(Note: In Eq. [255] the primes may be arbitrarily removed from x_1' and X', and in Eq. [256] they may be arbitrarily added to y_2 and Y. This makes the first of those two equations applicable to object-point O, and the second to image-point I'.)

Thus we have, by Eqs. [255] and [256],

$$x_1' = -27.461538 + \frac{10.5(5)}{1.276950} = 13.652053$$

$$y_2 = 84 - \frac{10.5(5)}{4.238514} = +71.613583.$$

We then have from Eq. [254], by numerical substitutions

$$a = \frac{-7(13.652053 + 15.461538)(71.613538 - 67) - (-0.619048)(13.652053)(71.613583)}{-67}$$

$$a = +4.999996$$

which closely checks the arbitrarily assigned value of a. As a further check, a can be calculated from Eq. [16], Part 1, as

$$a = \frac{D - S_1 - S_2}{2} \pm$$

$$\left[\left[\frac{D - S_1 - S_2}{2}\right]^2 - (D - S_1 - S_2)(f_1 + f_2) - \frac{(1 - M)^2}{M} f_1 f_2\right]^{1/2} \quad [257]$$

or

$$a = 23.750000 \pm \sqrt{351.562490}$$
$$a = 23.750000 \pm 18.750000$$
$$a = 42.500000 \text{ or } 5.000000,$$

which again checks the arbitrarily assigned value of a. Among other numerous check possibilities from Sec. 7 of Part 1, Eq. [13] may be written

$$M' = \frac{f_1 f_2}{(f_1 + f_2 - a)x_1' + f_1(f_2 - a)} \quad [258]$$

from which

$$M' = \frac{4.238514(1.276950)}{(4.238514 + 1.276950 - 5.0)(13.652053) + (4.238514)(1.276950 - 5.0)}$$
$$M' = -0.619048,$$

which exactly checks the value calculated following Table 16.

In this example, D has been one of the arbitrary parameters. If, instead, D_M or D_M' had been arbitrary, Eq. [252] could still have been used by first calculating D from

$$D = D_M - \Delta_c y \text{ (see Eq. [181], Part 2)} \quad [259]$$

or

$$D = D_M' + \Delta_c x \text{ (see Eq. [245])}. \quad [260]$$

Similarly, for D'

$$D' = D_M - \Delta_c x \text{ (see Eq. [181], Part 2)} \quad [261]$$

or

$$D' = D_M' + \Delta_c y \text{ (see Eq. [245])}, \quad [262]$$

which may be useful when Eq. [252] is expressed in terms of D', X', and Y'.

The problem discussed in this example has Eq. [252] as an important part of its solution. Readers must expect that every problem is likely to have a different algebraic solution, which in many cases can be reached or approached with the aid of the tables, schematic

diagrams, and equations that are developed in the text and appendices of the three parts of this paper; or may be reached by contriving new solutions in the general manner of the solutions found herein.

Example 15. In Example 14, the signed numerical values of X, Y, and Y' were the initially given parameters of Table 16. It proved possible through calculation of additional parameters to identify the problem as belonging to Group IV of Table 15. If the sign of Y' is merely reversed, without any change in the values of X and Y, will the problem then automatically belong to Group III_a (as Table 15 appears to suggest), and what additional restrictions, if any, will result? It is assumed that the sign of f (or F) is held unchanged.†

This question is readily examined by expressing X' as in (III-5), Table 1, Part 1, with f and Y' substituted *in signet form* (see footnote to Sec. 28, *RCA Rev.* 45, No. 1, p. 136); thus,

$$X' = \frac{fY'}{f - Y'} = \frac{(f > 0)(Y' < 0)}{(f > 0) - (Y' < 0)} = \frac{(fY') < 0}{(f - Y') > 0} = (X' < 0). \qquad [263]$$

This means that for *any* $(Y' < 0)$ and *any* $(f > 0)$ there will *always* be an $(X' < 0)$, as required by Group III_a.

Furthermore, from (V-5), Table 1, $M' = (f - Y')/f$ to produce

$$M' = \frac{(f > 0) - (Y' < 0)}{(f > 0)} = \frac{(f - Y') > 0}{(f > 0)} = (M' > 0) \qquad [264]$$

as required by Group III_a; and since in the numerator of Ineq. [264]

$$(f > 0) - (Y' < 0) = (f > 0) + [-(Y' < 0)]$$
$$= (f > 0) + (-Y' > 0) = [(f - Y') > 0] > (f > 0), \qquad [265]$$

then (for, the sum of two positive quantities—both > 0—is always greater than either), it is also true that $(M' > 0) > 1$, as required by Group III_a.

The sign of $M = Y/X$, since the signs of Y and X remain opposed, remains $(M < 0)$, as required by group III_a.

The signs of $\Delta_c x$ and $\Delta_c y$ are all that remain to be determined, and it will suffice to determine the sign of only one of them. By substituting the appropriate signets into Eq. [43], Part 1,

$$\Delta_c y = -f(M' - M) = -(f > 0)[(M' > 0) - (M < 0)]$$

† If the signs of *both* Y' and f (or F) of Group IV are changed, to what group will the signet-substitution process (about to be; discussed), show the problem of Example 14 to be converted? See Appendix 9.

or

$$\Delta_c y = -(f > 0)[(M' > 0) + (-M > 0)] = (\Delta_c y < 0) \qquad [265A]$$

as required by Group III_a. Then by Eq. [232]

$$\Delta_c x = \frac{\Delta_c y}{MM'} = \frac{(\Delta_c y < 0)}{(M < 0)(M' > 0)} = (\Delta_c x > 0) \qquad [266]$$

as required by Group III_a.

This completes demonstration that the mere sign reversal of Y' automatically changes the problem of Group IV to Group III_a in Table 15. Additional restrictions are that the parameters in columns (3)', (7)', (10), and (11) will have to be recomputed even though there is no change in the *numerical* values of the parameters in columns (2)*, (4)*, and (5)'. Changes in the numerical values of these latter three parameters will, however, also require recomputation of the parameters in columns (1), and (6)*. The statements of this paragraph can be checked along the lines of the calculations that follow Table 16.

Using the general technique of signet substitution that has just been illustrated, readers should prove in detail that reversal of the signs of *both* X' and Y in Group IV will convert the problem to the form of Group $VIII_b$ in every respect. Derive, for Groups IV and $VIII_b$, the nn.-metrics that relate X directly to Y and X' directly to Y' in signet form.† Readers should also study the result of reversing the signs of *both* X' and Y' in Group IV, and should also prove that reversal of the sign of X' only, in Group IV, in all respects converts the problem to the form of Group III_b.

Example 16. (Note: This example of the Gaussian design of an auxiliary optical system is expressed in terms of, among other things, an auxiliary magnification $(M > 0) > 1$. It is natural to inquire about the effect of changing this magnification to $(M > 0) < 1$. The effect of such a change on the procedure of simple Gaussian design of the auxiliary optical system is not trivial, and is the subject of Example 19, Sec. 30A).

An optical system (see Fig. 17) illuminates a transparent object, such as a photographic negative, by projecting the image of a light source through it and focusing the source image upon the plane of the entrance pupil EP of a projection lens. Without making any other changes whatsoever, it is required to place an auxiliary lens (or a separated pair of auxiliary lenses) between the transparency and the projection lens, in order to produce a small increase in the size of the final projected image of the transparency without shifting

† Example: A required nn.-metric for Group IV is $(Y > 0) \lesseqgtr (-X > 0)$ or $(-Y < 0) \gtreqless (X < 0)$ or any others among their numerous equivalents.

the location of the transparency's image along the axis. The light source must not shift its location along the axis and must at the same time remain focused on the plane of the entrance pupil of the projection lens, which also must not move to a new location respecting the transparency.

Without going into details, a little consideration of these requirements in relation to Table 15 indicates that the system to be designed probably corresponds to Group I_b, and that there is therefore apparently no *fundamental* reason to expect serious difficulty in working out a Gaussian design that will meet the stated requirements.

The distance of the projection lens entrance pupil from the transparency is 12.5 inches, and the auxiliary lens is to magnify the image of the transparency by the factor $(M > 0) = +1.020$. The projection lens itself, as well as the original magnification of the transparency's image, is in no sense a part of the problem.

For the auxiliary lens or separated pair, let the transparency be the real object O of Fig. 17, and the source image that is projected onto the entrance pupil of the projection lens be the virtual object O′. The conditions of the problem require the virtual image I (of O), that is produced by the auxiliary lens or separated pair, to be *in the plane of* O. They also require the real image I′ (of O′) to be *in the plane of* O′.

Thus, the distance from O to I, which is OI, is

D = OI = Zero,

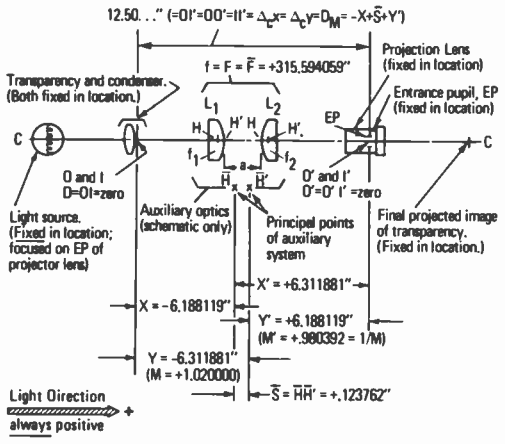


Fig. 17—Diagram of auxiliary optical system problem of Sec. 29A, Example 16. For schedule of solutions, see Table 17. Schematic only—not to scale.

and the distance from O′ to I′, which is O′I′, is

$D' = \text{O'I'} = \text{Zero.}$

From this circumstance, and from Eq. [181], Part 2,

$\text{OI'} = \text{OO'} = \text{II'} = \Delta_c x = \Delta_c y$, or $\Delta_c y/\Delta_c x = 1$.

Therefore, from Eq. [232],

$\Delta_c y/\Delta_c x = MM' = 1$, or $M = 1/M'$,

a result that is also confirmed by the established relationships of optical system photometry (see Refs. 23–28, Part 2). Furthermore, Eq. [181] shows that since

$D' = D = \text{Zero} = \text{OI} = \text{O'I'}$

then

$\Delta_c x = \Delta_c y = \text{OI'} = D_M$

which is the distance from the transparency to the projection lens entrance pupil. Thus

$\Delta_c x = \Delta_c y = 12.5$ inches

and Eq. [43], Part 1, enables us to calculate the focal length of the auxiliary system as

$$f = \frac{-\Delta_c y}{M' - M}; \qquad [267]$$

or on taking note that $M' = 1/M$

$$f = \frac{-M\Delta_c y}{1 - M^2}$$

giving

$$f = \frac{-1.02(12.5)}{1 - (1.02)^2}$$

$f = +315.594059$ inches.

Since D, M, and f are now all known quantities, and since by (I-8) Table 1, Part 1,

$(D - S) = -f(1 - M)^2/M$

we may now express the separation of the two principal points of the auxiliary as

$$S = \frac{f(1 - M)^2}{M} + D \qquad [268]$$

or

$$S = \frac{+315.594059(1 - 1.02)^2}{1.02} + \text{Zero}$$

$$S = +0.123762 \text{ inch.}$$

These results are readily checked by use of the equation (see Sec. 4, Part 1, p. 441)

$$M_{1,2} = 1 - \frac{D - S}{2f} \pm \left[\left[1 - \frac{D - S}{2f}\right]^2 - 1\right]^{1/2} \qquad [269]$$

from which, given that $D = D' = 0$, $S = +0.123762$, and $f = +315.594059$,

$$M_{1,2} = +1.000196 \pm 0.019804$$

or

$$M_1 = 1.020000$$
$$M_2 = 0.980392 \ (= 1/M_1).$$

This provides an indirect but exact confirmation of the correctness of the values of S, D, and f. The justification of this method of check is left to the reader.

We may also write, from (III-10) Table 1, Part 1,

$$X = \frac{D - S}{M - 1} = \frac{-.123762}{1.02 - 1} = -6.188119 \text{ inches,} \qquad [269a]$$

and from (IV-10) Table 1,

$$Y' = \frac{(D' - S)M'}{M' - 1} = \frac{-0.123762(0.980392)}{0.980392 - 1} = +6.188119 \text{ inches,} \qquad [269b]$$

and since $D' = 0$, and $Y' + S = X'$ (because $Y' - X' = D' - S$, from (I-1) Table 1), and since also $X' - X = \Delta_c x$ by Eq. [172], it follows that

$$-X + S + Y' = \Delta_c x = 2(6.188119) + 0.123762$$
$$= 12.500000 \text{ inches.}$$

This confirms the values of S, D, D', X, and Y', and it shows that because $X = -Y'$, the principal points of the lens, or separated

pair, are centered exactly midway between the transparency and the projection lens entrance pupil.

The auxiliary may be a single lens, with a focal length $f = +315.594059$ inches and principal plane separation $S = 0.123762$ inch. All planolenses and most ordinary double-convex or double-concave lenses have positive principal plane separations. Planolenses, whether convex or concave, have positive values of S which are *independent* of the focal length of the lens, but they do depend on the thickness and the refractive index of the glass. For any planolens, $S = d(n - 1)/n$, where $d > 0$ is the axial thickness of the lens and $n > 1$ is the refractive index of the glass. Also, one of the principal points *always* coincides with the vertex of the spherical surface of a plano-lens, the other one being inside the glass.

These simplifying relationships make plano-lenses very easy and inviting to include in the planning stages of an optical system, where preliminary design work is based on Gaussian equations. In cases where S is negative or very small or zero, problems arise of which Example 19, Sec. 30A, is an instance. The theory of Sec. 30 is very useful in such cases.

The auxiliary may also be a pair of separated lenses, specified as f_1, S_1, and f_2, S_2 with an optical separation a. Their combined focal length must still be $\overline{F} = +315.594059$ inches, and the separation of the principal points of the combination must still be $\overline{S} = 0.123762$ inch. The additional design latitude that becomes available with the use of two lenses, instead of only one, will now be briefly discussed.

By Eq. [45], Part 1,

$$\overline{F} = \frac{f_1 f_2}{f_1 + f_2 - a} \qquad [270]$$

from which, by setting $f_2 = kf_1$, k being an arbitrary constant which might be termed a "shape factor,"

$$kf_1^2 - \overline{F}(1 + k)f_1 + \overline{F}a = 0 \qquad [270a]$$

or

$$f_1 = \frac{\overline{F}(1 + k)}{2k} \pm \left[\frac{\overline{F}^2(1 + k)^2}{4k^2} - \frac{\overline{F}a}{k}\right]^{1/2} \quad (f_2 = kf_1). \qquad [271]$$

Equation [271] represents the full gamut of real possibilities open to f_1 (and also to f_2) when values are assigned to a, $\overline{F}$, and k, subject, of course, to the requirement that the discriminant of Eq. [270a],

$\overline{F}^2(1 + k)^2 - 4k\overline{F}a$ must be ≥ 0.

Another factor that is important in this example is the separation $\overline{S}$ between the principal points of the two-lens auxiliary system. By Eq. [30], Part 1, this is expressed as

$$\overline{S} = \frac{-a^2}{f_1 + f_2 - a} + S_1 + S_2 \text{ or } S_1 + S_2 = \overline{S} + \frac{a^2}{f_1(1 + k) - a}.$$

On then substituting the value of f_1 from Eq. [271]

$$S_1 + S_2 = \overline{S} + \frac{a^2}{\dfrac{\overline{F}(1 + k)^2}{2k} \pm \left[\dfrac{\overline{F}^2(1 + k)^4}{4k^2} - \dfrac{\overline{F}a(1 + k)^2}{k}\right]^{1/2} - a}$$

which may be reduced further to

$$S_1 + S_2 = \overline{S} + \frac{2ka^2}{(1 + k)[\overline{F}(1 + k) \pm \{\overline{F}^2(1 + k)^2 - 4k\overline{F}a\}^{1/2}] - 2ka} \quad [272]$$

in which, since it is necessary for Eq. [272] to be conformable to Eq. [271], it is still necessary that the discriminant of Eq. [270a],

$\overline{F}^2(1 + k)^2 - 4k\overline{F}a$ shall be ≥ 0.

Only one simple case of the two-lens auxiliary will be considered in detail, that in which $k = 1$, or $f_2 = f_1$, with the lens separation a ranging over the reasonable interval $0 \leq a \leq 15$ inches. In this case Eq. [272] readily reduces to the form

$$S_1 + S_2 = \overline{S} + \frac{a^2}{2[\overline{F} \pm (\overline{F}^2 - \overline{F}a)^{1/2}] - a} \quad (k = 1 \text{ or } f_2 = f_1). \quad [273]$$

Equation [271] correspondingly reduces to

$$f_1 = f_2 = [\overline{F} \pm (\overline{F}^2 - \overline{F}a)^{1/2}] \quad (k = 1 \text{ or } f_2 = f_1). \quad [274]$$

In order to make the solution complete it is necessary to solve for both x_1 and y_2' and to show that their values satisfy the equation (see Fig. 12. Note: In *this* example I′ happens to coincide with O′)

$$-x_1 + S_1 + a + S_2 + y_2' = \text{OI}' = D_M = \Delta_c x. \quad [275]$$

Equation [275] clearly follows from inspection of Fig. 12. The value of x_1 is obtained by solving Eq. [13], Part 1, as

$$x_1 = \frac{f_1 f_2 - Mf_1(f_2 - a)}{M(f_1 + f_2 - a)}. \quad [276]$$

By applying the replacement transformation of Sec. 6, Part 1, to Eq. [13], noting that in this example $M' = 1/M$

$$M' = \frac{-(f_1 + f_2 - a)y_2' + f_2(f_1 - a)}{f_1 f_2} \qquad [277]$$

from which

$$y_2' = \frac{M' f_1 f_2 - f_2(f_1 - a)}{-(f_1 + f_2 - a)} . \qquad [278]$$

Thus the two-lens auxiliary system will consist of two identical double-convex or plano-convex lenses of focal lengths f_1 and f_2 (with $f_2 = f_1$), optically separated a distance a, with the center point of distance a coinciding with the center point of the distance D_M = OI′, between the transparency and the projection lens entrance pupil. If plano-convex lenses are mounted with their convex surfaces facing each other, touching them together makes $a = 0$. This property of plano-lenses is mentioned further in Sec. 30. If readers wish to study the cases in which $k \neq 1$, they will find it very helpful to have a clear understanding of Sec. 30 and of Example 19, Sec. 30A.

Table 17 presents assumed values of the parameter a and the values of other parameters that are calculated from the indicated equations. In the interest of keeping the two lenses from becoming too thick, it is desirable to keep $S_1 + S_2$ in the range 0.125 inch $\leq (S_1 + S_2) \leq$ 0.375 inch.† In the case of the example discussed here, this requirement is easily met for all values of a that are possible within the limit that is imposed by the 12.5-inch separation between the transparency and the projection lens entrance pupil.

30. A Brief Introduction to *Any* Single, Thick, Spherical Lens In Air or Vacuum, Treated As a System of Two Quasi-Separated Lenses.

A single, solid, thick lens element L with a focal length F, is a physical piece of homogeneous refracting material of index of refraction n, bounded by two opposed spherical refracting surfaces of arbitrary curvature. The two centers of curvature K_1 and K_2 (or K and K′) determine a unique right line called the axis; C-C in Fig. 18. The surfaces intersect the axis in two vertex points V_1 and V_2 (or V and V′) of the figure. Their separation is the axial lens thickness d. If an object O on C-C is imaged paraxially at I, the distances $\bar{x}_1$ (from V_1 to O) and $\bar{y}_2$ (from V_2 to I) are vertex-conjugate dis-

† If the chosen index of refraction is $n_1 = n_2 = n$, and the two identical lenses are assumed plano-convex, the thickness of each lens is $d_1 = d_2 = d = n\,(S_1 + S_2)/[2(n - 1)]$.

Table 17—Calculated values of the variable parameters $f_1 = f_2$, $S_1 + S_2$, $x_1 = -y_2'$, and check $OI' = \Delta_c x$, as a is varied arbitrarily between zero and 15 inches. The calculations depend on the given constants: $M = +1.020$, $\overline{F} = +315.5940594$ inches, and $\overline{S} = +0.123762$ inch. All calculated values are in inches. Numbers in square brackets identify the equations by which the associated quantities are calculated.

Item No.	a (Assumd)	[274] $f_1 = f_2$	[273] $(S_1 + S_2)$†	[276] $x_1 = -y_2'$	[275] $OI' = \Delta_c x$
1	0	+631.188118	+0.123762	−6.188119	+12.500000
2	+0.3	631.038082	0.123833	−6.038083	12.499999
3	0.7	630.837924	0.124151	−5.837924	12.499999
4	1.0	630.687721	0.124555	−5.687722	12.499999
5	4.0	629.181740	0.136517	−4.181741	12.499999
6	8.0	627.162443	0.175113	−2.162443	12.499999
7	12.0	625.129972	0.240054	−0.129973	12.500000
8	15.0	623.596817	0.306363	+1.403182	12.499999

† If the chosen index of refraction is $n_1 = n_2 = n$, and the two identical lenses are assumed plano-convex, the thickness of each lens is $d_1 = d_2 = d = n(S_1 + S_2)/[2(n - 1)]$.

tances;†† these contrast with the Gaussian conjugate distances X and Y, that are measured from the principal points $\overline{H}$ and $\overline{H}'$ of the lens. The vertex-conjugate distances x_1 and y_2 can be very important both practically and theoretically, in working with single, thick lenses (see examples 17 and 18).

An imagined plane P, normal to C-C, is considered to divide the lens L into two parts or components, L_1 and L_2 (or L and L'). The dividing plane may cross the axis anywhere inside the lens material, but for convenience it is arbitrarily assumed to be located midway between the vertex points V_1 and V_2. When considered in this way, any simple thick lens may be regarded as a centered system of two quasi-separated, plano-spherical component lenses with their hypothetical plano-surfaces in contact, coincident or "submerged" in the glass, but with their adjacent principal points $\widehat{H}_1'$ and H_2 spaced apart by an optical separation $\bar{a}$. (To avoid confusion, it is important to note that these "adjacent" principal points *are not* the principal points of lens L. The matter will be considered fully in the second paragraph below.) The individual focal lengths $\widehat{f}_1$ and $\widehat{f}_2$ (or $\widehat{f}$ and $\widehat{f}'$), of the assumed plano-components $\widehat{L}_1$ and $\widehat{L}_2$, respectively, are independent of the interior location of plane P. This is because in any positive or negative thick plano-lens, whose index of refraction and finite radius of curvature are given, the focal length is *never* a function of its thickness. Thus, in Eq. [287] below, $\overline{F}$ is independent of d when r_1 (or r_2) $= \infty$. The optical separation

†† When a single thick lens is resolved into plano-components, it is helpful if some of the symbols in equations that pertain to such components are distinguished by an "arc" written over them. This practice is followed as sparingly as possible.

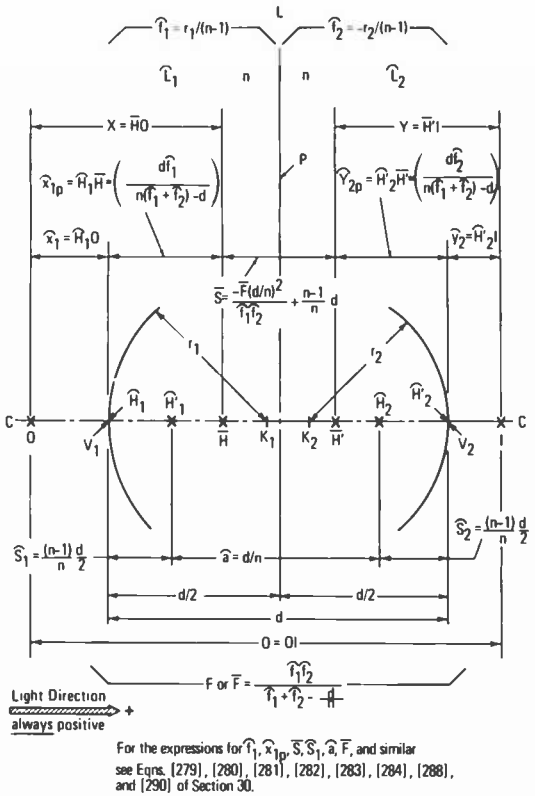


Fig. 18—Diagram of a thick spherical lens, in illustration of its treatment in Sec. 30 as a system of two quasi-separated lenses. Schematic only—not to scale.

$\hat{H}'_1\hat{H}_2 = \hat{a}$, of the component plano-lenses, into which the plane P is assumed to divide any thick lens, is also independent of the plane P's location (see Eqs. [281] and [282], below).

By regarding the simple thick lens in the above manner, many of the relationships derived earlier for two separated lenses may be advantageously applied in dealing with the Gaussian problems of any thick lens (see Part 1, Secs. 5 through 9).

The focal length of a plano-spherical lens in air or vacuum is positive when the spherical surface is outwardly convex and negative when the surface is outwardly concave. The light direction, however, is always positive, and for lens L it is directed from V_1 toward V_2. Accordingly, the well-known Gaussian formula for any plano-spherical lens in air or vacuum expresses the radius of curvature of the first or anterior surface of any single thick spherical lens as

$$r_1 = (n - 1)\widehat{f}_1 \qquad [279]$$

and that of the second or posterior surface as

$$r_2 = -(n - 1)\widehat{f}_2. \qquad [280]$$

A radius is always measured from the spherical surface toward its center of curvature, and is positive in the direction of the light and negative opposed. The thickness of the lens-component does not appear in either of these two equations; the equations are accordingly independent of lens-component thicknesses.

The overall axial thickness of lens L is d, each plano-component being arbitrarily assigned the thickness $d/2$. Well-known Gaussian formulas then express the principal point separation for each plano-lens as

$$\widehat{S}_1 = \widehat{S}_2 = \widehat{H}_1\widehat{H}_1' = V_1\widehat{H}_1' = \widehat{H}_2\widehat{H}_2' = \widehat{H}_2V_2 = (n - 1)d/2n \qquad [281]$$

where $\widehat{H}_1$ *always* coincides with V_1, and $\widehat{H}_2'$ *always* coincides with V_2. The optical separation $\widehat{a}$ between the component plano-lenses can now be expressed as

$$\widehat{a} = \widehat{H}_1'\widehat{H}_2 = d - (\widehat{S}_1 + \widehat{S}_2) = d - 2(n - 1)d/2n = d/n. \qquad [282]$$

Although the locations of the points $\widehat{H}_1'$ and $\widehat{H}_2$ are locations of convenience that depend on the arbitrary or assumed location of dividing plane P, their separation $\widehat{a}$ is fixed, and depends only on d and n. But the two points $\overline{H}$ and $\overline{H}'$ are the principal points of lens L, whose *fixed* locations are given by Eqs. [32] and [33], Part 1, as the principal distances (in this case vertex-principal distances)

$$\widehat{H}_1\overline{H} = V_1\overline{H} = \widehat{x}_{1p} = \frac{\widehat{f}_1 d/n}{\widehat{f}_1 + \widehat{f}_2 - d/n} \quad (d/n = \widehat{a}) \qquad [283]$$

and

$$\widehat{H}_2'\overline{H}' = V_2\overline{H}' = \widehat{y}_{2p} = \frac{-\widehat{f}_2 d/n}{\widehat{f}_1 + \widehat{f}_2 - d/n} \quad (d/n = \widehat{a}). \qquad [284]$$

Furthermore, by suitable substitutions from Eq. [45], Part 1,

$$\widehat{H}_1\overline{H} = V_1\overline{H} = \widehat{x}_{1p} = \frac{\overline{F}d/n^{\dagger}}{\widehat{f}_2} \quad (d/n = \widehat{a}) \qquad [285]$$

and

† This paper has not always been consistent in its choice of a symbol for the focal length of a system of two *separated* lenses; f, F, and $\overline{F}$ have been variously used, sometimes with applicable subscripted numeral(s) or letter(s), or with a distinguishing prime. In Part 3 of the paper the writer favors $\overline{F}$ as the basic symbol for this quantity.

$$\bar{H}_2'\bar{H}' = V_2\bar{H}' = y_{2p} = \frac{-\bar{F}d/n}{\bar{f}_1} \quad (d/n = \bar{a}). \tag{286}$$

where $\bar{F}$ is the overall focal length of the *combined* plano-spherical components $\bar{L}_1$ and $\bar{L}_2$ (whose hypothetical plano-surfaces are in contact, coincident, or "submerged" in the glass).

In the preceding paragraph, distances x_{1p} and y_{2p} (see Sec. 9, Part 1) are called principal distances, for the first time. Moreover, when the two separated lenses are plano-spherical, with their plano-surfaces facing each other (whether in contact or not), their principal distances may be written as x_{1p} and y_{2p} and termed vertex-principal distances (note that in Eqs. [283] through [286] the "air-space" between the two plano-surfaces *must* be zero).

By equating either the two expressions for $\bar{x}_{1p}$ (Eqs. [283] and [285]) or, alternatively, those for $\bar{y}_{2p}$ (eqs. [284] and [286]) and solving for $1/\bar{F}$, we readily obtain the usual expression that relates $\bar{F}$, $\bar{f}_1$, $\bar{f}_2$, d, and n of a thick lens in air or vacuum, as

$$\frac{1}{\bar{F}} = (n - 1)\left[\frac{1}{r_1} - \frac{1}{r_2} + \frac{n - 1}{n} \cdot \frac{d}{r_1 r_2}\right] \tag{287}$$

where $r_1 = (n - 1)\bar{f}_1$ and $r_2 = -(n - 1)\bar{f}_2$ (see Eqs. [279] and [280]).

If, from Eq. [45], Part 1, by allowing $\bar{a} = d/n$, as in Eq. [282],

$$\bar{F} = \frac{\bar{f}_1\bar{f}_2}{\bar{f}_1 + \bar{f}_2 - d/n} \quad (d/n = \bar{a}) \tag{288}$$

and if, from Eq. [30], by additionally allowing $\bar{S}_1 + \bar{S}_2 = (n - 1)d/n$ in accordance with Eq. [281], we write

$$\bar{S} = \frac{-(d/n)^2}{\bar{f}_1 + \bar{f}_2 - d/n} + \frac{n - 1}{n} d^{\dagger} \quad (d/n = \bar{a}), \tag{289}$$

it immediately follows by inspection of Eqs. [288] and [289], that

$$\bar{S} = \frac{-\bar{F}(d/n)^2}{f_1\bar{f}_2} + \frac{n - 1}{n} d \quad (d/n = \bar{a}) \tag{290}$$

from which

$$\bar{S} = \frac{-\bar{F}d^2 + n(n - 1)f_1 f_2 d}{n^2\bar{f}_1\bar{f}_2} \tag{290a}$$

† The remarks of the foregoing footnote substantially apply to the symbol s, S, or $\bar{S}$ for the separation of the two principal points of a system of two *separated* lenses. When the k-th lens of a system (in Sec. 30A, Example 19) is treated as a quasi-separated system of two lenses, its focal length can be symbolized as either f_k or as $\bar{F}_k$, with an asterisk added if it is an "adjusted" focal length. Similarly, the principal point separation for such a system may appear as S_k or as $\bar{S}_k$.

and

$$\overline{F}d^2 - n(n-1)\widehat{f}_1\widehat{f}_2 d + n^2\widehat{f}_1\widehat{f}_2\overline{S} = 0. \quad [291]$$

The solution of Eq. [288] for $\widehat{f}_2$ is

$$\widehat{f}_2 = \frac{(\widehat{f}_1 - d/n)\overline{F}}{\widehat{f}_1 - \overline{F}} \quad (d/n = \widehat{a}), \quad [292]$$

which, on substitution into Eq. [291], produces

$$(\widehat{f}_1 - \overline{F})d^2 + (n-1)\widehat{f}_1 d^2 - n\widehat{f}_1\overline{S}d + n^2\widehat{f}_1^2\overline{S} - n(n-1)\widehat{f}_1^2 d = 0 \quad [293]$$

or

$$(n\widehat{f}_1 - \overline{F})d^2 - n[\widehat{f}_1\overline{S} + (n-1)\widehat{f}_1^2]d + n^2\widehat{f}_1^2\overline{S} = 0. \quad [294]$$

The solution for d, for example, is

$$d = \frac{n[\widehat{f}_1\overline{S} + (n-1)\widehat{f}_1^2]}{2(n\widehat{f}_1 - \overline{F})} \pm \left[\frac{n^2[\widehat{f}_1\overline{S} + (n-1)\widehat{f}_1^2]^2}{[2(n\widehat{f}_1 - \overline{F})]^2} - \frac{n^2\widehat{f}_1^2\overline{S}}{n\widehat{f}_1 - \overline{F}}\right]^{1/2\dagger} \quad [295]$$

for *all* arbitrary values of n, $\widehat{f}_1$, $\overline{F}$, and $\overline{S}$, but with the double proviso that (on dividing the discriminant of Eq. [294] by the always positive $n^2\widehat{f}_1^2$):

$$[\overline{S} + (n-1)\widehat{f}_1]^2 - 4\overline{S}(n\widehat{f}_1 - \overline{F}) \geq 0 \quad [296]$$

and

$$n\widehat{f}_1 - \overline{F} \neq 0. \quad [297]$$

Equation [293] may be restated and solved for any variable to suit the reader's purpose; variables may always be removed and replaced with new ones by suitable substitution. The possibilities are far too numerous to consider here. Readers may wish to study the comments in Conrady's *Applied Optics and Optical Design*, p. 64 and pp. 455–456, regarding the "bending" of both thin and thick lenses, and to consider how the Conrady procedures may be related to Eq. [293] and any of its solutions. There is no conflict between the equations developed in this paper and Conrady's sign convention.

---

† In a problem that involves a different set of Gaussian parameters, Appendix 11 derives a much simpler expression for the center thickness d of a simple lens (see Eq. [438]).

30A. Representative Applications

Example 17: Identify Some of the Conditions That Must Be Met, If a Single, Thick Lens Is to Have Coincident Principal Points

This can be done by analyzing the solution of Eq. [293] for $\overline{S}$, which is

$$\overline{S} = d\left[\frac{\widehat{f}_1^2(n^2 - n) - (n\widehat{f}_1 - \overline{F})d}{n\widehat{f}_1(n\widehat{f}_1 - d)}\right]. \qquad [298]$$

The principal points are coincident when $\overline{S} = 0$. This will *always* be true of thin lenses, for they are characterized by $d = 0$. In a thick lens $d \neq 0$, and in this case $\overline{S} = 0$ only when the numerator in Eq. [298] vanishes, or when

$$\widehat{f}_1^2(n^2 - n) - (n\widehat{f}_1 - \overline{F})d = 0 \quad (\overline{S} = 0). \qquad [299]$$

This will be true provided

$$\widehat{f}_1 = \frac{nd}{2(n^2 - n)} \pm \left[\frac{(nd)^2}{4(n^2 - n)^2} - \frac{\overline{F}d}{n^2 - n}\right]^{1/2} \quad (\overline{S} = 0). \qquad [300]$$

The requirement for this solution to be real is that the discriminant of Eq. [299]

$n^2d^2 - 4(n^2 - n)\overline{F}d$ shall be ≥ 0.

Since n and d are *always* positive, this requirement for the discriminant may be written as

$$\overline{F} \leq \frac{nd}{4(n - 1)} \quad \begin{matrix}(d > 0);\\ (n > 1);\end{matrix} \qquad [301]$$

or as

$$\overline{F} = \left[\frac{gnd}{4(n - 1)} \leq \frac{nd}{4(n - 1)}\right] \quad \begin{matrix}(d > 0);\\ (n > 1);\\ [-\infty < (g \neq 0) \leq 1].\end{matrix} \qquad [302]$$

By making suitable substitutions into Eq. [300] from Eqs. [279] and [302] and reducing

$$r_1 = (n - 1)\widehat{f}_1 = \frac{d}{2} \pm \frac{d}{2}(1 - g)^{1/2} \quad (\overline{S} = 0). \qquad [303]$$

Furthermore, from Eqs. [280] and [292] we may write

$$r_2 = -(n - 1)\widehat{f}_2 = \frac{-(n - 1)(\widehat{f}_1 - d/n)\overline{F}}{\widehat{f}_1 - \overline{F}} \qquad [304]$$

and by introducing into Eq. [304] the expressions for $\widehat{f}_1$ from Eq. [279] and for $\overline{F}$ from Eq. [302]

$$r_2 = \frac{-(n-1)\left[\dfrac{r_1}{n-1} - \dfrac{d}{n}\right]\dfrac{gnd}{4(n-1)}}{\dfrac{r_1}{n-1} - \dfrac{gnd}{4(n-1)}}$$

which reduces to

$$r_2 = -(n-1)\widehat{f}_2 = \frac{[(n-1)d - nr_1]gd}{4r_1 - gnd} \quad (\overline{S} = 0). \qquad [305]$$

By further simplifying Eqs. [303] and [305], their difference yields

$$r_1 - r_2 = d \quad (\overline{S} = 0). \qquad [306]$$

This result requires the radii r_1 and r_2 to originate at a common center of curvature for the two surfaces of the lens.

By substituting $\widehat{f}_1 = \dfrac{r_1}{n-1}$ and $\widehat{f}_2 = \dfrac{-r_2}{n-1}$ from Eqs. [279] and [280] and also by substituting $d = r_1 - r_2$ from Eq. [306] into Eqs. [283] and [284], the expressions for the vertex principal distances result:

$$\widehat{x}_{1p} = r_1 \quad (\overline{S} = 0) \qquad [307]$$

and

$$\widehat{y}_{2p} = r_2 \quad (\overline{S} = 0). \qquad [308]$$

Equations [307] and [308] place the first and second principal points of the lens at the common center of curvature of the anterior and posterior surfaces of the lens, respectively, (see Sec. 9, Part 1).

It has been proved that in order for a single thick spherical lens element to have coincident principal points:

(1) The lens must be concentric, or the centers of curvature of both lens surfaces must be coincident;

(2) The principal points will be located coincidentally at the common center of curvature of the lens surfaces.

This conclusion is independent both of the index of refraction and of the sign of the focal length of the lens. By substitution into Eq. [287] from Eq. [306] it is easily shown that the lens will be double convex (not necessarily equi-convex) when of positive focal length, and will be concavo-convex (or convexo-concave) when of negative

focal length. Readers will also find it stimulating to study the relationship between this example and Eq. [295].

Example 18: Discuss Theoretically At Least Two Possible Gaussian Procedures For the Nondestructive Measurement of the Index of Refraction of the Material of a Single Thick Spherical Lens Element

Accurate measurement of the indices of refraction of an optical glass is a painstaking procedure that ordinarily involves spectroscopic methods and equipment to enable measurements to be made at several wavelengths in the spectrum. The glass sample on which the measurements are made is usually in the form of a carefully made prism.

It cannot be expected that measurements of index of refraction, that depend on the Gaussian properties of a spherical lens into which the glass has already been fabricated, can be made with the accuracy that is possible when the spectroscopic method is used; but it may sometimes be desirable, useful, or even imperative that such a measurement be made, without modifying the lens in any way.

Possible Procedure No. 1: Equation [287] for any single thick lens is quadratic in n. It can be solved for the index of refraction in the form

$$n = \frac{\bar{F}(r_2 - r_1 + 2d) + r_1 r_2}{2\bar{F}(r_2 - r_1 + d)} \pm \left[\frac{[\bar{F}(r_2 - r_1 + 2d) + r_1 r_2]^2}{[2\bar{F}(r_2 - r_1 + d)]^2} - \frac{d}{r_2 - r_1 + d} \right]^{1/2}. \qquad [309]$$

Measurements of d, r_1, and r_2 may usually be made with reasonable accuracy with a micrometer and test glasses or other good spherometric techniques. However, the paraxial focal length $\bar{F}$, in light of a definite wavelength or color such as provided by a laser, generally must be measured indirectly, usually by some magnification method that nearly always involves numerous uncertainties and inaccuracies. In order to avoid the effect of lens aberrations on judgment of the paraxial image location, the lens must be stopped down. Unfortunately the judgment of image location then remains uncertain due to the increased depth of focus of the image. There is a possibility of overcoming this difficulty by using a pinhole as a test-object and employing the Foucault knife-edge technique to locate the true paraxial image. The magnification can then be measured by moving the pinhole a measured, signed distance t_o trans-

verse to the lens axis and measuring the corresponding, signed transverse motion t_i of its image. The magnification is then $M = t_i/t_o$. Keeping the lens location *fixed,* the pinhole is moved along the lens axis and the magnification M' of the new paraxial image is measured. The axial separation of the two pinhole locations is $\Delta_c x$ and that of the two paraxial images is $\Delta_c y$. By Eqs. [43] and [42], Part 1, the focal length is then calculated as

$$\overline{F} = \frac{-\Delta_c y}{M' - M} = \frac{-MM'\Delta_c x}{M' - M}. \qquad [310]$$

Substitution of the measured and calculated values of the four variables r_1, r_2, d, and $\overline{F}$ into Eq. [309] allows n to be calculated. Note that these four quantities are fixed, or constant, for any given lens, and except for measurement errors they are ideally independent of the kinds of measurements from which they are derived.

It is possible, under certain conditions, for the two values of n that are determined by Eq. [309] to be numerically close together, in which case the equation offers no clue to which of them is the true index of the measured lens. This always present, potential ambiguity of Procedure No. 1, together with the anticipated freedom from it of Possible Procedure No. 2 now to be developed, is discussed fully in Appendix 10.

Possible Procedure No. 2: (Discussion of this procedure is fully illustrated by Fig. 18.) It is clearly desirable to avoid measurement of the magnification of any image produced by the lens as a step in the measurement of its index of refraction.

That the possibility of doing so may exist becomes apparent when the simple, thick spherical lens is regarded as a quasi-separated system of two centered plano-spherical component lenses, with plano-surfaces hypothetically coincident, in contact, or "submerged" in the glass in the manner of the approach that is taken to such lenses by Sec. 30. For then, any of the equations that characterize a system of two separated lenses may be applied to learn if the magnification measurements can be eliminated.

By applying the equations for systems of two separated lenses, such as Eqs. [7] through [22] and Eqs. [63] and [64], Part 1, to "Possible Procedure No. 2," it is seen that the vertex-conjugate object-distances $\tilde{x}_1$ are measured from $\widehat{H}_1$ of lens-component $\widehat{L}_1$ and that the vertex-conjugate image-distances $\tilde{y}_2$ are measured from $\widehat{H}'_2$ of lens-component $\widehat{L}_2$. It is assumed that these distances are measured in light of a definite wavelength or color, such as provided by a laser.

In the case of any simple lens, $\widehat{H}_1$ and $\widehat{H}'_2$ coincide, respectively,

with the vertex points V_1 and V_2, as pointed out at the beginning of Sec. 30. This ensures that in the case of a simple lens, $\bar{x}_1$ and $\bar{y}_2$ can *always* be measured from the vertices of the lens surfaces, which differs importantly from measuring X and Y from possibly inaccessible and usually unknown principal points $\overline{H}$ and $\overline{H}'$ elsewhere on the lens axis.

Also among these equations, Eq. [19], Part 1, might be a basis for the desired Procedure No. 2, because magnification M does not appear in it. For these several reasons the possibility of the existence of a more practical alternative to Procedure No. 1 as a means of measuring the index of refraction of a simple lens appears favorable.

Before proceeding further it will be useful to identify some quantities and their equivalents that are involved in working with equations that relate to a single thick, simple lens. These are collected in Table 18, where each quantity is identified by an item number. The equivalents assigned to the symbols stem largely from Eqs. [279] through [282], which are derived in Sec. 30.

In the following discussions of Procedure No. 2 of this example, all item numbers refer to Table 18, unless otherwise referenced. Since the equations pertain to components of lens L, f_1 and f_2 have been changed to $\widehat{f}_1$ and $\widehat{f}_2$ to indicate that fact, etc.

By Eq. [19], Part 1,

$$D - \widehat{S}_1 - \widehat{S}_2 = \widehat{a} - \bar{x}_1 + \frac{\widehat{f}_2[(\widehat{f}_1 - \widehat{a})\bar{x}_1 - \widehat{f}_1\widehat{a}]}{(\widehat{f}_1 + \widehat{f}_2 - \widehat{a})\bar{x}_1 + \widehat{f}_1(\widehat{f}_2 - \widehat{a})} \qquad [311]$$

or, by substituting a suitable equivalent to D from item 15

$$-\bar{y}_2 + \frac{\widehat{f}_2[(\widehat{f}_1 - \widehat{a})\bar{x}_1 - \widehat{f}_1\widehat{a}]}{(\widehat{f}_1 + \widehat{f}_2 - \widehat{a})\bar{x}_1 + \widehat{f}_1(\widehat{f}_2 - \widehat{a})} = 0. \qquad [312]$$

(Special NOTE: Equation [312] could also have been derived (without benefit of Eq. [19]) by eliminating M between Eq. [13] and its replacement transformation (see Sec. 6, Part 1) and then solving for $\bar{y}_2$ and reducing. This useful shortcut procedure once again demonstrates and confirms a possible application of the replacement transformation. Details are left to the reader.)

On further substitution into Eq. [312] of equivalents to $\widehat{f}_1$, $\widehat{f}_2$, and $\widehat{a}$, from items 3, 6, and 12

$$-\bar{y}_2 + \frac{-kr_2\left[\left(kr_1 - \dfrac{kd}{1+k}\right)\bar{x}_1 - kr_1\dfrac{kd}{(1+k)}\right]}{\left(kr_1 - kr_2 - \dfrac{kd}{1+k}\right)\bar{x}_1 + kr_1\left(-kr_2 - \dfrac{kd}{1+k}\right)} = 0 \qquad [313]$$

Table 18—Some quantities and their equivalents that are involved in working with equations that relate to a single lens. See also Fig. 18.

Item No.	Description	Symbol and its Equivalent
1	Index of refraction	n
2	Radius of curvature; anterior surface	r_1
3	Focal length; anterior plano-component $\bar{L}_1$	$\bar{f}_1 = r_1/(n-1) = kr_1$; [$k = 1/(n-1)$; (see Eq. [279]) $\therefore n = (1/k) + 1$]
4	Axial thickness; anterior plano component $\bar{L}_1$	$\bar{d}_1 = d/2$
5	Radius of curvature; posterior surface	r_2
6	Focal length; posterior plano-component $\bar{L}_2$	$\bar{f}_2 = -r_2/(n-1) = -kr_2$; [$k = 1/(n-1)$; (see Eq. [280]) see item 3]
7	Axial thickness; posterior plano-component $\bar{L}_2$	$\bar{d}_2 = d/2$
8	Principal point separation; anterior plano-component $\bar{L}_1$	$\bar{S}_1 = (n-1)\bar{d}_1/n = \bar{d}_1/(1+k)$; [$k = 1/(n-1)$; (see Eq. [281] and item 4) $\therefore 1 + k = n/(n-1)$; see item 3]
9	Principal point separation; posterior plano-component $\bar{L}_2$	$\bar{S}_2 = (n-1)\bar{d}_2/n = \bar{d}_2/(1+k)$; [see item 8]
10	Overall axial thickness of simple lens L	$d = \bar{d}_1 + \bar{d}_2 = \bar{S}_1 + \bar{a} + \bar{S}_2$
11	Sum of component principal point separations	$\bar{S}_1 + \bar{S}_2 = (\bar{d}_1 + \bar{d}_2)/(1+k) = d/(1+k) = (n-1)d/n$; [see items 7 and 8]
12	Optical separation of plano-components $\bar{L}_1$ and $\bar{L}_2$	$\bar{a} = d/n = kd/(1+k)$; [$n = (1/k) + 1$; (see Eq. [282]) $\therefore n = (1+k)/k$; see item 3]
13	Vertex-conjugate object-distance V_1O	$\bar{x}_1$
14	Vertex-conjugate image-distance V_2I	$\bar{y}_2$
15	Overall object-to-image distance	$D = -\bar{x}_1 + d + \bar{y}_2 = -\bar{x}_1 + \bar{S}_1 + \bar{a} + \bar{S}_2 + \bar{y}_2$; [see items 13, 10, 14, 8, 12, and 9]

which expands to

$$-\bar{y}_2(r_1 - r_2 - d)\bar{x}_1 - k\bar{y}_2(r_1 - r_2)\bar{x}_1 + kr_1\bar{y}_2(r_2 + d) + k^2r_1r_2\bar{y}_2 - kr_2(r_1 - d)\bar{x}_1 - k^2r_1r_2\bar{x}_1 + k^2r_1r_2d = 0 \quad [314]$$

and factors to the quadratic in k

$$k^2r_1r_2(\bar{y}_2 - \bar{x}_1 + d) - k[\bar{x}_1\bar{y}_2(r_1 - r_2) + r_1r_2(\bar{x}_1 - \bar{y}_2) - (\bar{x}_1r_2 + \bar{y}_2r_1)d] - \bar{x}_1\bar{y}_2(r_1 - r_2 - d) = 0. \quad [315]$$

The solution for k is

$$k = \frac{\bar{x}_1\bar{y}_2(r_1 - r_2) + r_1r_2(x_1 - y_2) - (x_1r_2 + \bar{y}_2r_1)d}{-2r_1r_2(\bar{x}_1 - \bar{y}_2 - d)} \pm \left[\frac{[\bar{x}_1\bar{y}_2(r_1 - r_2) + r_1r_2(\bar{x}_1 - \bar{y}_2) - (\bar{x}_1r_2 + \bar{y}_2r_1)d]^2}{[2r_1r_2(\bar{x}_1 - \bar{y}_2 - d)]^2} - \frac{\bar{x}_1\bar{y}_2(r_1 - r_2 - d)}{r_1r_2(\bar{x}_1 - \bar{y}_2 - d)}\right]^{1/2}. \quad [316]$$

Then, by item 3 the solution for n is

$$n = \frac{1}{k} + 1 = \frac{1 + k}{k} \text{ (see Table 18, items 3 and 12).} \quad [317]$$

Equations [316] and [317] make possible the direct calculation of the index of refraction n of a simple lens from directly measured values of r_1, r_2, d, $\bar{x}_1$, and $\bar{y}_2$, without measurement of any magnification and without calculation of any focal length. The longitudinal axial location of the paraxial image I should be determined by the Foucault knife-edge technique. Negative lenses and lenses of very short focal length will present special problems, and the procedure affords ample opportunity for development of technique.

Readers will wish to test Eqs. [309], [316], and [317] by actual numerical substitution, and for that purpose they will be interested in the following (for a further comparative discussion of System Nos. 1 and 2, see Appendix 10):

Hypothetical Problem: It is necessary to duplicate a certain important simple lens as accurately as possible, but no specimen of the lens is available for measurements. Accurate computer readout records of certain design data at numerous specified wavelengths can be located, but, strange to say, there are no records of the indices of refraction. The records that are available include (for sodium-d light)

1. $\bar{F} = +3.9370 \ldots$ u (where u is an arbitrary unit)
2. $d = +0.750 \ldots$ u

3. $r_1 = +50.34011553$ u
4. $r_2 = -2.211106823$ u
5. $\widehat{x}_1 = -4.59320 \ldots$ u
6. $\widehat{y}_2 = +17.69591504$ u

The value of n obtained by calculating k and n from Eqs. [316] and [317] should be verified by also calculating n from Eq. [309]. They will be found to differ by less than one part in a billion (10^9, *not* the British 10^{12}). The other two solutions for n are negative, and their ratio is about 13.45.

Readers will also be interested in estimating the relative accuracies of the two procedures, in case n were calculated from actual measurements performed on an accurately made specimen of the lens that is described by items 2, 3, and 4 in the preceding list. This can be done by estimating the probable minimum inaccuracies that occur in the actual measurements, and by then using them in the combinations that will produce the greatest inaccuracies in the calculated values of n that result from each procedure. Detailed investigation of this point is left to the interested reader.

Example 19: Develop the Gaussian Design Procedures For Auxiliary Optical Systems Similar to the Earlier Example 16 of Sec. 29 But In Which $k = f_2/f_1 \neq 1$. (See Eqs. [272], [273], and [274])

I. Sample Specifications, and Preliminary Stage of Solution

The cited Example 16 of Sec. 29 is radically changed in its solution when a certain apparently trivial change is made in the system's specification. The specification change consists only in replacing $M = +1.020$ by its reciprocal, $M = 1/1.020 = +0.980392$ (which was the M' of Example 16). By then preliminarily applying the procedure of Example 16 in every detail, but without going beyond the end of the paragraph that contains Eq. [269b], the *new* conditions, on which the solution of Example 19 must be based, are found to be

$f = -315.594060$ inches (usually designated $\overline{F}$ for a separated system of two lenses)
$S = -0.123762$ inch (usually designated $\overline{S}$ for a separated system of two lenses)
$X = -6.311881$ inches
$Y' = +6.311881$ inches
(X and Y' calculated by (III-8) and (IV-8), Table 1, Part 1.

Compared to Example 16, f and S merely have their signs reversed, the relation $X = -Y'$ remaining unchanged, but X and Y' assume new values without changing sign. Also, the relationship $-X + S + Y' = \Delta_c x = +12.500$ inches remains unchanged.

II. General Analysis, and Application of Equations For the Primary Stage of the Solution

It is easy to see that the "new conditions" exclude the type of primary-stage solution that was convenient in Example 16, where the auxiliary could consist of two identical plano-lenses symmetrically arranged in the system. In that system $S_1 = S_2 > 0$, and the optical separation of the lenses was $a \geqslant 0$. By slightly restating Eq. [273] of Example 16

$$(S_1 + S_2) - \overline{S} = \frac{a^2}{2[\overline{F} \pm (\overline{F}^2 - \overline{F}a)^{1/2}] - a} \quad \begin{matrix}(k = 1, \text{ or } f_2 = f_1, \\ a \geqslant 0, \overline{F} > 0, \overline{S} > 0).\end{matrix} \qquad [318]$$

This equation was made a basis for the solution of Example 16, but because the application of the preliminary procedures of Example 16 (through Eq. [268] of that example) to Example 19 demonstrates a reversal of the signs of $\overline{F}$ and $\overline{S}$, the assumption that S_1, S_2, and a may all remain positive requires examination.

If we write (as representative of *the new conditions* of Example 19) the hypothetical function (based on Eq. [318]) as

$$\phi(a) = \frac{a^2}{2[\overline{F} \pm (\overline{F}^2 - \overline{F}a)^{1/2}] - a} \quad \begin{matrix}(a \geqslant 0, \overline{F} < 0, \overline{S} < 0, \\ k = 1, \text{ or } f_2 = f_1),\end{matrix}$$

it may be shown routinely that

1. $\phi(a > 0) < 0$,
2. $\phi(0) = 0$,

and that

3. $\lim_{a \to 0^+} \phi(a) = \phi(0)$.

Conclusions 1 and 2 demonstrate that Eq. [318] holds true *under the new conditions* of Example 19, provided only that $(S_1 + S_2) < 0$ (since $\overline{S} < 0$); while conclusions 2 and 3 demonstrate that $\phi(a)$ is continuous from the right at $(a \geqslant 0) = 0$. Therefore, S_1 and S_2 may *not* both remain positive.

In the solutions about to be developed, the condition $k = 1$ or $f_2 = f_1$ that defines the derivation of Eq. [318] from Eq. [272] will be

merely a special or limiting case. The solutions will apply generally, in the case of any value of $k = f_2/f_1$ for which the values of f_1, expressed by Eq. [271], are real. The condition $k = +1$ will apply only when the discriminant of Eq. [324] below, as expressed by Eq. [326], vanishes.

Only one approach to solution of the problem, that defined by the new conditions, will be examined and completed here. Possible modifications of this approach, that will require readers to apply some effort and originality to their analyses and solutions, will then be briefly indicated and defined.

We begin with Eq. [45], Part 1, which states

$$\overline{F} = \frac{f_1 f_2}{f_1 + f_2 - a} \qquad [319]$$

from which

$$\frac{1}{f_1 + f_2 - a} = \frac{\overline{F}}{f_1 f_2} \qquad [320]$$

and, from which,

$$f_2 = \frac{\overline{F}(f_1 - a)}{f_1 - \overline{F}}. \qquad [321]$$

From Eq. [30], Part 1,

$$\overline{S} = \frac{-a^2}{f_1 + f_2 - a} + S_1 + S_2 \qquad [322]$$

or, by substitution into Eq. [322] from Eq. [320],

$$\overline{S} = \frac{-a^2\overline{F}}{f_1 f_2} + S_1 + S_2. \qquad [323]$$

On substituting the value of f_2 into Eq. [323] from Eq. [321], the former reduces to the quadratic in f_1,

$$f_1^2(\overline{S} - S_1 - S_2) + f_1 a[a - (\overline{S} - S_1 - S_2)] - a^2\overline{F} = 0, \qquad [324]$$

which has the conjugate solutions

$$f_1 = \frac{a[a - (\overline{S} - S_1 - S_2)]}{-2(\overline{S} - S_1 - S_2)} \pm \left[\frac{a^2[a - (\overline{S} - S_1 - S_2)]^2}{4(\overline{S} - S_1 - S_2)^2} + \frac{a^2\overline{F}}{\overline{S} - S_1 - S_2}\right]^{1/2}. \qquad [325]$$

Equations [319] through [323] are all symmetrical in f_1 and f_2.

This means that the two values of f_1 that are represented by Eq. [325] can be substituted, respectively, for f_1 and f_2 in Eq. [321] and satisfy that equation. This is easily verified by arbitrary numerical substitutions into Eqs. [325] and [321]. The conjugate solutions of Eq. [324] are therefore the symmetrically related focal lengths f_1 and f_2 of *any* Gaussian system of two separated lenses. Only real solutions of Eq. [324] are of any interest, of course.

In many problems where the locations and focal lengths of two lenses are sought, Eq. [325] would complete the solution. This is not the case here, for the unknowns a and $(S_1 + S_2)$ are not evaluated. The optical separation a can be given an arbitrary value, but this cannot be done for $(S_1 + S_2)$. Once a is arbitrarily given, the following procedure can be used to find quasi-arbitrary values of $(S_1 + S_2)$ that will *always* result in real (not complex) solutions of Eq. [324].

The discriminant of Eq. [324] is

$$\partial_1 = a^2[a - \overline{S} + (S_1 + S_2)]^2 + 4a^2\overline{F}[\overline{S} - (S_1 + S_2)] \qquad [326]$$

or

$$\partial_1 = a^2[(S_1 + S_2)^2 + 2(a - \overline{S} - 2\overline{F})(S_1 + S_2) + (a - \overline{S})^2 + 4\overline{FS}], \qquad [327]$$

which itself happens, in this example, to have the general form of a quadratic in $(S_1 + S_2)$.

In order for the roots of Eq. [324] to be real, it is necessary that $\partial_1 = (\partial_1 \geqslant 0)$. Thus the variable $\partial_1 = (\partial_1 \geqslant 0)$ has the extremum, or lower limit, $\partial_1 = (\partial_{1X} = 0)$, corresponding to which the radicand of Eq. [325] vanishes, and Eq. [327] has two limiting solutions for $(S_1 + S_2)$. These are

$$(S_1 + S_2)_X^{\mp} = -(a - \overline{S} - 2\overline{F}) \pm [(a - \overline{S} - 2\overline{F})^2 - (a - \overline{S})^2 - 4\overline{FS}]^{1/2} \qquad [328]$$

to which the special condition $k = +1$, or $f_2 = f_1$ of Eq. [318], corresponds (for when $\delta_1 = (\delta_{1X} = 0)$, Eqs. [324] and [325] *always* yield $f_2 = f_1$).

Note: The limiting solutions of Eq. [327] are keyed to the mixed sign in the root on the right side of Eq. [328], as are the limiting derivatives expressed by Eq. [330], below. In the cases both of the limiting solutions and the limiting derivatives, the subscript X signifies that the discriminant $\partial_1 = (\partial_{1X} = 0)$.

Since ∂_1 can have an infinity of values $\partial_1 > (\partial_{1X} = 0)$ for which the conjugate solutions of Eq. [324] are real, there will be quasiarbitrary values $(S_1 + S_2) = (S_1 + S_2)_X^{\mp} + d(S_1 + S_2)$, which second

addend is a differential or increment of $(S_1 + S_2)$, that correspond to the positive values $\partial_1 = (\partial_{1X} = 0) + d\partial_1$, which second addend is a differential or increment of ∂_1. The problem now is to find the algebraic sign that the differential or increment $d(S_1 + S_2)$ must have. By differentiation of Eq. [327]

$$\frac{d\partial_1}{d(S_1 + S_2)} = 2a^2[(S_1 + S_2) + a - \overline{S} - 2\overline{F}]. \qquad [329]$$

Substitution of $(S_1 + S_2) = (S_1 + S_2)_{\overline{X}}^{\pm}$ into Eq. [329] results in two limiting values of the derivative, which are

$$\left[\frac{d\partial_1}{d(S_1 + S_2)}\right]_X^{\pm} = 2a^2[(S_1 + S_2)_{\overline{X}}^{\pm} + a - \overline{S} - 2\overline{F}]. \qquad [330]$$

Since the limits expressed by Eqs. [328] and [330] assume that the discriminant $\partial_1 = (\partial_1 \geqslant 0)$ is at the limit $\partial_1 = (\partial_{1X} = 0)$, any differential or increment $d\partial_1$ that is added to ∂_{1X} *must* be positive or the solution of Eq. [324] will be complex. Therefore, the sign of any differential or increment $d(S_1 + S_2)$ that is added to either of the limits $(S_1 + S_2)_{\overline{X}}^{\pm}$ *must* be that of the corresponding derivative

$$\left[\frac{d\partial_1}{d(S_1 + S_2)}\right]_X^{\pm}$$

when $(S_1 + S_2)_{\overline{X}}^{\pm}$ is substituted into Eq. [330].

Table 19 is arranged in 10 Roman-numeral columns, each with an identifying symbol and an identifying equation number by which the corresponding signed numerical values of the symbols in each column are calculated. Each of the 30 Arabic-numeral rows of the table identifies one trial, partial-solution of Example 19. The rows are arranged in six groups of five, each group being identified by a different assumed optical separation a (Column III) between lenses L_1 and L_2. Individual rows are further identified by the quasi-arbitrary sum $(S_1 + S_2)$ (Column VI) of the principal point separations, S_1 of L_1 and S_2 of L_2. The uppermost sum in Column VI of each group has been adjusted slightly, as needed, to ensure real values of f_1 and f_2. Rounding-off errors in computing Column IV make this necessary. The other sums $S_1 + S_2$ in Column VI of any group represent successive incremental additions of $0 < d(S_1 + S_2) \leqslant +0.002$ inch.† Column IV also gives the second solutions $(S_1 + S_2)_{\overline{X}}^{\mp}$ of Eq. [328], which range between -1262.1 and -1262.9 inches; they can be investigated at the pleasure of the reader.

† *Carefully* review the above discussions that follow Eqs. [323] through [330].

Table 19—Collected preliminary-stage solutions (Columns I and II) and trial primary-stage solutions (Columns III through X), of Example 19 (see text following paragraph of Eq. [330]). Columns VII through X refer to lenses (L_1) and (L_2). All columns (except Col. V, which is a derivative) are in inches. The trial secondary-stage solutions collected in Table 20 essentially complete solution of Example 19. Item No. 18 judged best primary stage solution.

Col. No.:	I	II	III	IV		V	VI	VII	VIII	IX	X
Eq. No.:	[267]	[268]	[None]	[328]		[330]	[None]	[325]	[325]	[248]	[249]
Item No.	$\bar{F}$	$\bar{S}$	Arbitrary a	$(S_1 + S_2)_X^+$	$(S_1 + S_2)_X^-$	$\left[\frac{d\partial_1}{d(S_1 + S_2)}\right]_X^+$	Quasi-Arbitrary $(S_1 + S_2)$	$f_1^{+\backslash} = f_1(L_1)$	$f_1^{-\backslash} = f_2(L_2)$	$x_{1p}(L_1)$	$y_{2p}(L_2)$
1	−315.594060	−0.123762	−0.200	−0.123794	−1262.1	+50.479047	−0.123793	−549.637073	−740.885507	−0.085194	+0.114837
2	−315.594060	−0.123762	−0.200	−0.123794	−1262.1	50.479047	−0.122000	+96.638226	−74.136790	−0.851383	−0.653145
3	−315.594060	−0.123762	−0.200	−0.123794	−1262.1	50.479047	−0.120000	63.378239	−52.945597	−1.192145	−0.995907
4	−315.594060	−0.123762	−0.200	−0.123794	−1262.1	50.479047	−0.118000	50.298924	−43.556890	−1.449112	−1.254874
5	−315.594060	−0.123762	−0.200	−0.123794	−1262.1	50.479047	−0.116000	42.880717	−37.927406	−1.664201	−1.471963
6	−315.594060	−0.123762	−0.040	−0.123763	−1262.4	+2.019674	−0.123763	−432.485642	−1167.554358	−0.010812	+0.029189
7	−315.594060	−0.123762	−0.040	−0.123763	−1262.4	2.019674	−0.122000	+17.368208	−16.500149	−0.765070	−0.726832
8	−315.594060	−0.123762	−0.040	−0.123763	−1262.4	2.019674	−0.120000	11.779761	−11.394455	−1.107886	−1.071648
9	−315.594060	−0.123762	−0.040	−0.123763	−1262.4	2.019674	−0.118000	9.480930	−9.243248	−1.365728	−1.331490
10	−315.594060	−0.123762	−0.040	−0.123763	−1262.4	2.019674	−0.116000	8.149111	−7.982978	−1.581335	−1.549097

11	−315.594060	−0.123762	−0.010	−0.123762	−1262.5	+0.126236	−0.123761	+234.545551	−134.555551	−0.023455	−0.013456
12	−315.594060	−0.123762	−0.010	−0.123762	−1262.5	0.126236	−0.122000	4.255595	−4.208842	−0.749836	−0.741598
13	−315.594060	−0.123762	−0.010	−0.123762	−1262.5	0.126236	−0.120000	2.904679	−2.888098	−1.092740	−1.086502
14	−315.594060	−0.123762	−0.010	−0.123762	−1262.5	0.126236	−0.118000	2.344014	−2.336659	−1.350621	−1.346383
15	−315.594060	−0.123762	−0.010	−0.123762	−1262.5	0.126236	−0.116000	2.017846	−2.014963	−1.566252	−1.564014
16	−315.594060	−0.123762	+0.010	−0.123762	−1262.5	+0.126240	−0.123761	+234.558261	−134.548261	+0.023456	+0.013455
17	−315.594060	−0.123762	0.010	−0.123762	−1262.5	0.126240	−0.122000	4.265662	−4.198909	0.751610	0.739848
18	−315.594060	−0.123762	0.010	−0.123762	−1262.5	0.126240	−0.120000	2.914725	−2.878143	1.096520	1.082758
19	−315.594060	−0.123762	0.010	−0.123762	−1262.5	0.126240	−0.118000	2.354051	−2.326696	1.356404	1.340642
20	−315.594060	−0.123762	0.010	−0.123762	−1262.5	0.126240	−0.116000	2.027878	−2.004995	1.574039	1.556277
21	−315.594060	−0.123762	+0.040	−0.123763	−1262.6	+2.019930	−0.123761	+1870.058549	−270.018549	+0.046751	+0.006750
22	−315.594060	−0.123762	0.040	−0.123763	−1262.6	2.019930	−0.122000	17.409281	−16.461222	0.766879	0.725117
23	−315.594060	−0.123762	0.040	−0.123763	−1262.6	2.019930	−0.120000	11.820495	−11.355189	1.111718	1.067596
24	−315.594060	−0.123762	0.040	−0.123763	−1262.6	2.019930	−0.118000	9.521523	−9.203842	1.371575	1.325813
25	−315.594060	−0.123762	0.040	−0.123763	−1262.6	2.019930	−0.116000	8.189622	−7.943489	1.589196	1.541434
26	−315.594060	−0.123762	+0.200	−0.123794	−1262.9	+50.511047	−0.123793	−550.796099	−739.326482	+0.085373	−0.114596
27	−315.594060	−0.123762	0.200	−0.123794	−1262.9	50.511047	−0.122000	+96.864848	−73.963373	0.853379	+0.651617
28	−315.594060	−0.123762	0.200	−0.123794	−1262.9	50.511047	−0.120000	63.596518	−52.763875	1.196250	0.992488
29	−315.594060	−0.123762	0.200	−0.123794	−1262.9	50.511047	−0.118000	50.513715	−43.371681	1.455300	1.249538
30	−315.594060	−0.123762	0.200	−0.123794	−1262.9	50.511047	−0.116000	43.093469	−37.740158	1.672458	1.464696

An inspection of the table and a review of the development in the text of each equation cited at the tops of the columns will serve to clarify both. Computation (or assumption) of the quantities in any row proceeds in logical order from Col. I through Col. X.

III. Review of the Preliminary and the Primary Stages of the Solution

The work to this point is characterized as being in two stages: the preliminary stage and the primary stage of the Gaussian design. In the preliminary stage, calculations were performed directly on the given data of Example 19. Columns I and II, Table 19, are the partial result of that stage. The primary stage roughed out a collection of 30 trial systems, in terms (in Columns VII and VIII) of the focal lengths f_1 and f_2 of the two lenses L_1 and L_2, (respectively identified as the $f_1^{\dagger\vee}$ and $f_1^{-\vee}$ values of Eq. [325]), as well as in terms (in Columns IX and X) of the respective principal distances $x_{1p}{}^{\dagger}$ and $y_{2p}{}^{\dagger}$ from their appropriate principal points, to the auxiliary system's principal points $\overline{H}$ and $\overline{H}'$ (see Sec. 9, Part 1). This notion of principal distances will be expanded and clarified in the following secondary stage of the design work.

Specific focal lengths and lens separations have not been settled upon that will ensure that the auxiliary system will be restricted to the space between the transparency and the lens entrance pupil of the system that is to be modified by the auxiliary. With the sum $(S_1 + S_2)$ being negative, it is obvious that at least one of the two lenses L_1 and L_2 *must* have a negative principal point separation. Beyond this assertion, the values of S_1 and S_2 remain individually unknown. A major part of the work of the secondary stage will be concerned with the design of a simple lens L_2 with a negative principal point separation S_2. Readers should also consider the effect of alternatively setting $S_1 < 0$ in order that $(S_1 + S_2) < 0$.

IV. The Secondary Stage of the Solution

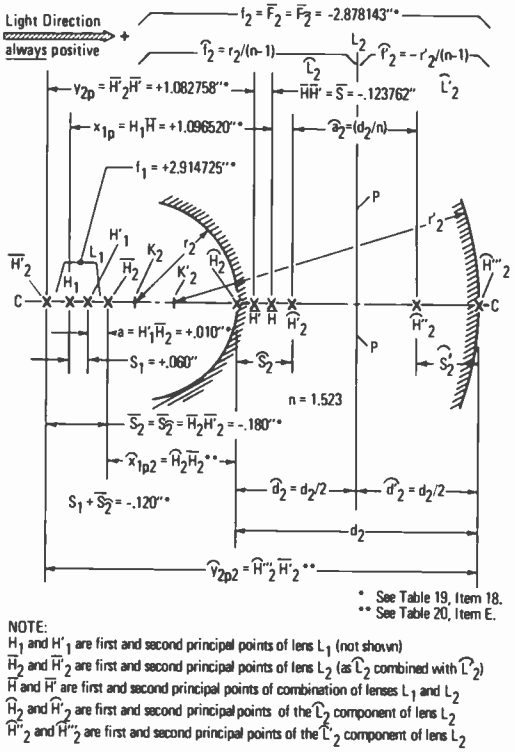
Ordinary simple lenses that are of common concern usually have their axial principal points either inside the lens, or not far from its surfaces if one or both of them are outside the lens. This has led to the habit and expectation of identifying the location of the principal points substantially with the location of the physical lens. It can sometimes happen, however, that the principal points will lie

† See footnote re: Steps 4 and 5, Table 21.

at a considerable distance from the physical lens, and this circumstance can lead to troublesome disorientation and confusion in working with such lenses if care is not taken to avoid it. The reason for this is that changes in paraxial ray directions that are assumed by the Gaussian equations take place (mathematically, not physically) at the transverse planes through the two principal points of the lenses or lens subsystems wherever those points may happen to be located along the axis. In contrast to this, the *actual* changes in ray directions take place, physically and sequentially, at each successive lens surface, where each change involves an initial ray direction, a refractive index, the radius of curvature of the surface where the change of direction is taking place, the thicknesses and separations of the individual lenses, etc. The Gaussian-lens-equation approach is much more general, because more abstract, and it "lumps" all of the data of actual, physical, surface-by-surface paraxial ray refraction in terms of much more general and inclusive parameters such as focal lengths, magnifications, object distances, system lengths, and the like. The advantages of the Gaussian approach have been made evident in the three parts of this paper where its use in the early stages of system planning has avoided much physical detail. In some cases the Gaussian abstractions and some of the concrete lens requirements are advantageously considered together, as will be done in the secondary stage of the solution of Example 19.

In Item 18 of Table 19, the linear quantities in Columns VII through X are small as a group and are considered likely to result in an auxiliary system of practical length that will readily fit between the transparency and the projection lens entrance pupil. The secondary stage of the work will accordingly be based on it without further justification. Lens L_1 ($f_1 = +2.914725$ inches) will be taken plano-convex, with its $S_1 = +0.060$ inch, merely because this is easy to do. In order to keep $(S_1 + S_2) = -0.120$ inch, S_2 of lens L_2 ($f_2 = -2.878143$ inches) must then be $S_2 = \overline{S}_2 = \overline{S}_2 = -0.180$ inch. It is clear from Example 17 that the principal points of a negative lens must lie at a considerable distance from the physical lens, if their separation is to be $S_2 = \overline{S}_2 = \overline{S}_2 = 0$, and that they may lie still further away if $\overline{S}_2 < 0$, as is now required.

The symbolism regarding lens L_2 in Fig. 19 corresponds roughly to the symbolism in Fig. 18 and to the footnotes that relate to Eqs. [285] and [289]. All symbols that pertain exclusively to lens L_2 of Fig. 19 or to its components $\widehat{L}_2$ and $\widehat{L}_2'$ are subscripted either solely or terminally with numeral 2 or with numeral $\widehat{2}$. The principal points $\overline{H}_2$ and $\overline{H}_2'$ of lens L_2 lie roughly an inch to the anterior side



* See Table 19, Item 18.
** See Table 20, Item E.

NOTE:
H_1 and H'_1 are first and second principal points of lens L_1 (not shown)
$\overline{H}_2$ and $\overline{H}'_2$ are first and second principal points of lens L_2 (as $\widehat{L}_2$ combined with $\widehat{L}'_2$)
$\overline{H}$ and $\overline{H}'$ are first and second principal points of combination of lenses L_1 and L_2
$\widehat{H}_2$ and $\widehat{H}'_2$ are first and second principal points of the $\widehat{L}_2$ component of lens L_2
$\widehat{H}'''_2$ and $\widehat{H}'''_2$ are first and second principal points of the $\widehat{L}'_2$ component of lens L_2

Fig. 19—Diagram illustrating unadjusted secondary stage solution of Example 19 of Sec. 30A. See also Tables 19 and 20 and Fig. 18. For adjustments, see Table 21. Schematic only—not to scale.

of the point $\widehat{H}_2$ of lens L_2 and of its anterior component $\widehat{L}_2$. The *NOTE* of Fig. 19, upon application of close inspection to the symbolism of the text and of Fig. 19, should allay any confusion. The symbols $\widehat{x}_{1p2}$ and $\widehat{y}_{2p2}$ of Fig. 19 are explained immediately following Eqs. [333] and [334]. The practice of using an "arc" (⌒) above a literal or numerical symbol has been followed as sparingly as possible in the interest of keeping the symbols simple.

Lens L_2 will have somewhat the form of Fig. 19. It will be meniscus, with its centers of curvature K_2 and K'_2 and its two principal points $\overline{H}_2$ and $\overline{H}'_2$ on the concave side of the lens. The lens will be considered as a system of two quasi-separated lenses of refractive index n and thickness $d_2 = \widehat{d}_2 + \widehat{d}'_2$, in the manner of Sec. 30 and Table 18. The optical separation will be $\widehat{a}_2 = d_2/n$. The two plano-component focal lengths will be $\widehat{f}_2$ and $\widehat{f}'_2$. The principal point separation of L_2 will be $S_2 = \overline{S}_2 = \overline{S}_{\widehat{2}} = \overline{H}_2\overline{H}'_2$. The focal length of L_2 will be $f_2 = \overline{F}_2 = \overline{F}_{\widehat{2}}$. If, in Eq. [293], we let $d = d_2$; $\overline{F} = \overline{F}_2$; $\overline{S} =$

$\overline{S}_2$; and $\widehat{f}_1 = \widehat{f}_2$ of the anterior $\widehat{L}_2$-component of lens L_2 (*not* the $\widehat{f}_2$ of Eq. [292]), the equation may easily be put in the standard form of the quadratic in $\widehat{f}_2$, as

$$[-n(n-1)d_2 + n^2\overline{S}_{\widehat{2}}]\widehat{f}_2^2 + [-n\overline{S}_{\widehat{2}}d_2 + nd_2^2]\widehat{f}_2 - \overline{F}_{\widehat{2}}d_2^2 = 0, \quad [331]$$

the conjugate solutions† of which are

$$\widehat{f}_2 = \frac{d_2(\overline{S}_{\widehat{2}} - d_2)}{2[n\overline{S}_{\widehat{2}} - (n-1)d_2]} \pm \left[\frac{d_2^2(\overline{S}_{\widehat{2}} - d_2)^2}{4[n\overline{S}_{\widehat{2}} - (n-1)d_2]^2} + \frac{\overline{F}_2 d_2^2}{n[n\overline{S}_{\widehat{2}} - (n-1)d_2]}\right]^{1/2}. \quad [332]$$

These are also expressed as the symmetrical solution set $\{\widehat{f}_2, \widehat{f}_2'\}$, where $\widehat{f}_2'$ is the focal length of the posterior plano-component L_2' of lens L_2.

The known or given quantities in Eq. [332] are

$\overline{S}_{\widehat{2}} = -0.180000$ inch
$n = 1.523000$
$\overline{F}_{\widehat{2}} = f_2 = -2.878143$ inches (Table 19, Item 18, Col. VIII).

Thickness d_2 is not known, but it may be considered to lie somewhere between the arbitrary but reasonable extremes of the interval $0 \leq d_2 \leq 1.50$ inches. By making d_2 the arbitrary parameter in Eq. [332], Table 20 (below) is constructed that relates d_2, $\widehat{f}_2$, $\widehat{f}_2'$, r_2, r_2', $\widehat{x}_{1p2}$, and $\widehat{y}_{2p2}$ (see text immediately following Eq. [334] below) of lens L_2. Equation [285] is readily reexpressed for present purposes as

$$\widehat{x}_{1p2} = \frac{\overline{F}_{\widehat{2}}d_2/n}{\widehat{f}_2'} \quad [333]$$

and Eq. [286] as

$$\widehat{y}_{2p2} = \frac{-\overline{F}_{\widehat{2}}d_2/n}{\widehat{f}_2} \quad [334]$$

where the numeral 2, that terminates the subscripts of vertex-principal distances $\widehat{x}_{1p2}$ and $\widehat{y}_{2p2}$, unmistakably identifies these quantities with the members of the quasi-separated pair L_2, L_2' of lens L_2. Note that it is important that plano-component $\widehat{L}_2$ be the negative component in lens L_2. If this is not done, lens L_2 may become physically

† See discussion of the meaning of the conjugate solutions of Eq. [324].

located anterior to lens L_1, although it will still be (mathematically or in the purely Gaussian sense) at optical separation $a = +0.010000$ inch (see Table 19, Item 18, Col. III) from L_1, and in that purely Gaussian sense posterior to L_1. This confusing point is worthy of very careful thought by the reader, for it is at such a stage in the work that the "troublesome disorientation and confusion" mentioned earlier can enter. The following two paragraphs should be helpful regarding the matter.

Assume that an optical system comprises a number of physical lenses spaced apart along a common axis or line of centers. The possibility then always exists of reversing the orientation of one of these lenses by rotating it 180° about, say, an axis normal to the system axis through one of the vertex points of the lens, and then moving the lens along the common axis until the axial or principal interval between its two principal points is restored to its original position along the system axis. The overall Gaussian or paraxial imagery of the system will be unaffected by such a reversal. This idea is more complicated than its simple statement suggests, and it invites careful consideration by readers. If this reversal and relocation cannot be accomplished without the "collision" of two physical lenses or without the interchange of their order along the axis, then the physical lens can have only one orientation in the system unless similar reversals of neighboring lenses can be performed to avoid such "collision" or order reversal. This "collision" or reversal of order is especially likely to occur if the principal points of the lens lie at a considerable distance *outside* the space included between the lens surfaces.

These physical considerations have been of importance in working out the solution to the secondary stage of Example 19 where they relate to the treatment, in Sec. 30, of a single thick lens as a system of two quasi-separated lenses. This treatment introduced physical considerations into the Gaussian equations for the first time in the form of the lens thickness d and the index of refraction n (see Eqs. [279] through [317], and Eqs. [331] through [334]). The matter is not considered further here, and additional study of details is left to the reader.

Table 20, which summarizes the results of the secondary stage of the solution of Example 19, is arranged in seven Arabic-numeral columns, numbered 1 through 7, with identifying symbols and with equation numbers by which the corresponding signed numerical values of the symbols in each column are calculated. Each of the nine capital-letter rows identifies one trial, secondary-stage solution of the example. Each row is also individually identified by an

Table 20—Plano-component focal lengths $\bar{f}_2$ and $\bar{f}_2'$, radii of curvature r_2 and r_2', and vertex principal distances $\bar{x}_{1p2}$ and $\bar{y}_{2p2}$, computed for lens L_2. The vertex principal distances locate the principal points $\bar{H}_2$ and $\bar{H}_2'$ of the quasi-separated system of L_2 with respect to H_2 of its plano-component $\bar{L}_2$ and with respect to H_2''' of its plano-component $\bar{L}_2'$. Thicknesses d_2 are assigned arbitrarily. Applicable equation numbers are given in the square brackets. All dimensions are in inches. (See Fig. 19).

Item No.	1 $d_2(L_2)$	2 [332] $\bar{f}_2(\bar{L}_2)$	3 [279] r_2	4 [332] $\bar{f}_2'(\bar{L}_2')$	5 [280] r_2'	6 [333] $\bar{x}_{1p2}$	7 [334] $\bar{y}_{2p2}$
A	+1.500	−1.140691	−0.596581	+3.521103	−1.841537	−0.805054	−2.485054
B	1.375	−1.104410	−0.577607	3.257033	−1.703428	−0.797798	−2.352798
C	1.250	−1.064114	−0.556532	2.990528	−1.564046	−0.789905	−2.219905
D	1.125	−1.019043	−0.532959	2.721187	−1.423181	−0.781280	−2.086280
E	1.000	−0.968223	−0.506380	2.448515	−1.280573	−0.771809	−1.951809
F	0.875	−0.910376	−0.476127	2.171881	−1.135894	−0.761350	−1.816350
G	0.750	−0.843792	−0.441303	1.890476	−0.988791	−0.749726	−1.679726
H	0.625	−0.766108	−0.400675	1.603234	−0.838491	−0.736708	−1.541708
J	0.500	−0.673960	−0.352481	1.308715	−0.684458	−0.722000	−1.402000

arbitrary value of parameter d_2 in Col. 1. The numbers in each row, Cols. 2 and 4, are conjugate solutions of Eq. [331]. The numbers in Table 19, Item 18, and in Table 20, Item E, together are considered to represent complete solution of Example 19 when they are supplemented by $X = -6.311881$ inches and $Y' = +6.311881$ inches from the preliminary stage of the work and when a is adjusted to the value $a^* = +0.009997$ inch, as shown in step 3 of the 7 steps of verification and adjustment in Table 21 (note also the adjusted values of f_2, x_{1p}, and y_{2p} of Cols. VIII, IX, and X, Table 19, Item 18 and steps 1, 4, and 5, Table 21). All quantities not found in Table 20 but which are involved in calculating the numbers in that table are given immediately following Eq. [332], above.

It is readily seen that by making some moderate changes, particularly in the values of a and $(S_1 + S_2)$ of Table 19, Cols. III and VI, and d_2 of Table 20, Col. 1, any number of other solutions are possible along the same general line of approach. Other lines of approach, some of which may be better or simpler, are always possible. A complete, adjusted, and summarized statement of the given and calculated quantities that represent the problem and its present solution follow the verification and adjustment procedures that are developed and grouped in Table 21. The complete summary statement appears as Table 22.

(Appendix 11 outlines how the literal quantities that basically define the simple, thick lens of Table 20, may be related in a simple manner.)

This essentially completes solution of the problem of Example 19. The arrangement of transparency, entrance pupil of the projection lens, and the auxiliary lenses that will slightly modify the magnification of the original optical system of Example 16, in the new way that has been required by Example 19, is illustrated by Fig. 20.

V. Verification and Adjustments

In general, the calculated numerical quantities in Table 20, Item E, were all considered "exact" as is the verification value $\underline{f}_2^* = \overline{F}_2^*$ of Table 21, step 1 (below), although it disagrees slightly with and will now replace for all further purposes f_2 of Table 19, Item 18, Col. VIII. The verification step 2 (Table 21) conflicts with this judgment but is brought (adjusted) into agreement with it by making a very slight adjustment of a to a new value a^* in step 3. This substantially restores the value of $\overline{F}$ in Table 19, Col. 1, on which all of that table was based. The value of $\underline{S}^* = -0.123760$ inch, which

Table 21—Calculation of "verification" and "adjusted" values of some of the quantities in Tables 19 and 20 to reduce the effects of the rounding-off errors on the system that represents a solution to Example 19. Quantities that appear in the calculations, but which are not otherwise identified, appear in Table 19, Item 18, also in Table 20, Item E, or in this table or following Eq. [332]. $S_1 = +0.060$ inch.

Step No.	Calculation of "Verification" and "Adjusted" Values
1.	By Eq. [288], calculate the verification (*and* adjusted) value of f_2, as $\underline{f}_2^* = \underline{F}_2^* = \dfrac{f_2 f_2'}{f_2 + f_2' - d_2/n} = -2.878146$ inches, and regard it as the exact or slightly adjusted value f_2^*, of f_2 in Table 19, It. 18, Col. VIII.
2.	By Eq. [270], calculate the verification value of $\overline{F}$, as $\underline{\overline{F}} = \dfrac{f_1 \overline{F}_2^*}{f_1 + \overline{F}_2^* - a} = -315.625272$ inches (considerably less than the exact $\overline{F}$ of Table 19, Col. I, and therefore calling for an adjusted value of a to be calculated in following step 3).
3.	By a solution of Eq. [319], calculate an adjusted (exact) value of a, as $a^* = \dfrac{(f_1 + \overline{F}_2^*)\overline{F} - f_1 \overline{F}_2^*}{\overline{F}} = +0.009997$ inch (to keep $\overline{F}$ at its original, exact -315.594060 inches of Table 19, Col. I; actually, within 0.004409 inch of original $\overline{F}$).
4.	By Eq. [248], calculate an adjusted (exact) value of x_{1p}, as $x_{1p}^* = \dfrac{a^* \overline{F}}{\overline{F}_2^*} = +1.096190$ inches (considerably less than the x_{1p} in Table 19, Item 18, Col. IX).†
5.	By Eq. [249], calculate the adjusted (exact) value of y_{2p}, as $y_{2p}^* = \dfrac{-a^* \overline{F}}{f_1} = +1.082433$ inches (considerably less than the y_{2p} in Table 19, Item 18, Col. X).†
6.	By Eqs. [323] and [285], calculate the verification value of $\overline{S}$, $\underline{\overline{S}} = \dfrac{-a^* x_{1p}^*}{f_1} + S_1 + \overline{S}_2 = -0.123760$ inch. (Note: $\overline{S}_2 = -0.180000$ inch; see given quantities that follow Eq. [332].)
7.	By Eqs. [323] and [286], calculate the verification value of $\overline{S}$, as $\underline{\overline{S}} = \dfrac{a^* y_{2p}^*}{\overline{F}_2^*} + S_1 + \overline{S}_2 = -0.123760$ inch (in precise agreement with the result of step 6, and differing so slightly from the $\overline{S} = -0.123762$ inch of Table 19, Col. II, that the latter value is accepted as "exact").

† These "considerable" changes in the principal distances x_{1p} and y_{2p} to the principal points $\overline{H}$ and $\overline{H}'$ of the system, are important only as they may indicate a change in $\overline{S} = \overline{HH}'$. Steps 6 and 7 show this change to be negligible.

Table 22—Summary of the symbolic quantities and their status as given (gvn), assumed (asd), calculated (cal), or adjusted (adj) values in Example 19 and its solution. See also Fig. 20.

Status and Item No.	Symbolic Quantities and Their Numerical Values
(gvn) 1	$M = +0.980392$
(gvn) 2	$OO' = \Delta_c x = X' - X = +12.500000$ in.
(gvn) 3	$II' = \Delta_c y = Y' - Y = +12.500000$ in.
(cal) 4	$M' = +1.020000$ (from $\Delta_c y/\Delta_c x = MM'$)
(cal) 5	$\overline{F} = -315.594060$ in.
(cal) 6	$\overline{S} = -0.123762$ in.
(cal) 7	$X = -6.311881$ in.
(cal) 8	$Y' = +6.311881$ in.
(cal) 9	$Y = MX = -6.188118$ in.
(cal) 10	$X' = Y'/M' = +6.188118$ in.
(gvn) 11	$D = OI =$ Zero
(gvn) 12	$D' = O'I' =$ Zero
(asd & adj) 13	$a^* = +0.009997$ in.
(asd) 14	$S_1 + S_2 = S_1 + \overline{S}_2 = -0.120000$ in.
(asd) 15	$S_1 = +0.060000$ in.
(cal) 16	$S_2 = \overline{S}_2 = -0.180000$ in.
(cal) 17	$f_1 = +2.914725$ in.
(gvn) 18	$n = +1.523000$
(asd) 19	$r_1 = \infty$
(cal) 20	$r_1' = -(n - 1)f_1 = -1.524401$ in.
(cal) 21	$d_1 = nS_1/(n - 1) = +0.174723$ in.
(cal & adj) 22	$f_2^* = \overline{F}_2^* = -2.878146$ in.
(cal & adj) 23	$x_{1p}^* = +1.096190$ in.
(cal & adj) 24	$y_{2p}^* = +1.082433$ in.
(cal) 25	$\overline{S}_2 = S_2 = -0.180000$ in.
(cal) 26	$f_2 = -0.968223$ in.
(gvn) 27	$n = +1.523000$
(cal) 28	$r_2 = (n - 1)f_2 = -0.506381$ in.
(cal) 29	$f_2' = +2.448515$ in.
(cal) 30	$r_2' = -(n - 1)f_2' = -1.280573$ in.
(cal) 31	$\bar{x}_{1p2} = -0.771809$ in.
(cal) 32	$y_{2p2} = -1.951810$ in.
(asd) 33	$d_2 = +1.000000$ in.
(cal) 34	$\bar{a}_2 = d_2/n = +0.656599$ in.
(cal) 35	$S_2 + S_2' = d_2 - \bar{a}_2 = +0.343401$ in.

very nearly verifies the value of $\overline{S}$ in Table 19, Col. II, can also be calculated from the adjusted values of x_{1p}^* and y_{2p}^* in steps 6 and 7. It is also easily shown, from Eqs. [289] and [290], that the values of S_1 and $S_2 = \overline{S}_{\bar{2}}$ are respectively dependent only on the construction of lenses L_1 and L_2, and are wholly independent of a or a^*.

As briefly as possible, the treatment of accumulated rounding-off errors (reduction of their effects, by adjustment) is indicated in steps 1 through 7 of Table 21, all referenced to Table 19, Item 18, and to Table 20, Item E. Note that in steps 2 through 7 the calculations involve some quantities that are calculated in the preceding step or

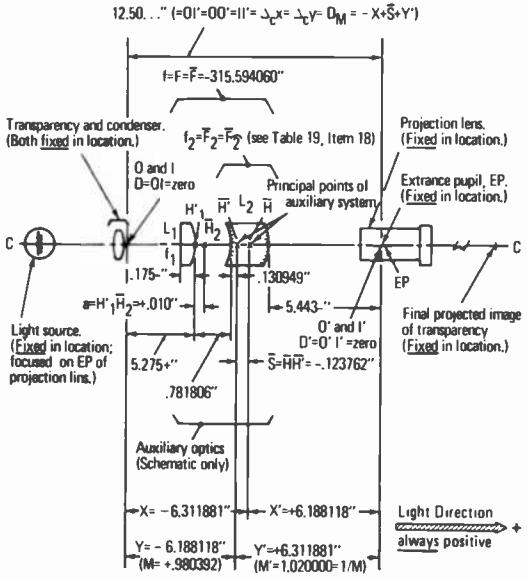


Fig. 20—Diagram of auxiliary optical system problem of Sec. 30A, Example 19. For schedules of solutions, see Table 19, especially Item 18, and Table 20, especially Item E. (See also Fig. 19.) For adjustments, see Table 21. Schematic only—not to scale.

steps. Calculation of an underlined symbol from Table 19, or Table 20, is a "verification" calculation. If the result is expected to be taken thereafter as the "exact," "true," or "adjusted" value of the symbol, it is "asterisked" in addition to being underlined; otherwise only the asterisk is retained and used, and the underline is dropped. The "adjusted" values of all other symbols are identified by asterisks.

Table 22 summarizes most of the numerical values of the symbols that pertain to Example 19 and its solution and contains the four adjusted (asterisked) values that were calculated in Table 21.

From the assembled data of Table 22 and the arrangement of Fig. 20, readers can readily verify that the distance of object O from the plane surface of lens L_1 approximates -5.101 inches, that the distance of image I′ from the convex surface of lens L_2 approximates $+5.443$ inches, and that the distance of the convex surface of lens L_2 from the plane surface of lens L_1 approximates $+1.957$ inches. Thus, the proposed auxiliary system, as here designed paraxially, will be easily accommodated in the $+12.500$-inch space between the transparency and the projection lens entrance pupil.

VI. Additional Procedures for Checking the Content of Table 22

At least three procedures are available for checking or determining the correctness of the results collected in Table 22.

1. Start with object O and calculate the location and magnification of each intermediate image (including the final image) of O, on a lens-by-lens basis, using the Gaussian equations of Table 1, Part 1. Multiply all of the intermediate magnifications together to get the value of the overall magnification M. Calculate the distance of the final image I from the initial object O by addition and subtraction of the appropriate conjugate distances, optical separations, principal point separations, lens thicknesses, air spacings, etc. Do the same, starting with object O′ and ending with image I′.
2. Perform the foregoing check by making a paraxial ray trace from O to I and from O′ to I′. This will base the check on radii of curvature and indices of refraction, instead of the Gaussian focal lengths. A performance of these traces proved to be an excellent check. Details of paraxial ray tracing are not a part of this paper, but may be found in many books on lens design and a few on geometrical optics.
3. From Eq. [13], Part 1, calculate the magnifications M (from object O to image I) and M' (from object O′ to image I′), as

$$M = \frac{f_1 f_2}{(f_1 + f_2 - a)x_1 + f_1(f_2 - a)} \qquad [335]$$

where, from Table 22,

$x_1 = X + x_{1p}^* = -6.311881 + 1.096190 = -5.215691$ inches
$x_1' = X' + x_{1p}^* = +6.188118 + 1.096190 = +7.284308$ inches
$f_1 = f_1 = +2.914725$ inches
$f_2 = f_2^* = -2.878146$ inches
$a = a^* = +0.009997$ inch
M(required) $= +0.980392$
M'(required) $= +1.020000$.

Also, from Eq. [19], Part 1, calculate the distances D (from object O to image I) and D' (from object O′ to image I′), as

$$D = a - x_1 + S_1 + S_2 + \frac{f_2[(f_1 - a)x_1 - f_1 a]}{(f_1 + f_2 - a)x_1 + f_1(f_2 - a)} \qquad [336]$$

where, also from Table 22,

$S_1 = \overline{S}_1 = +0.060000$ inch
$S_2 = \overline{S}_{\widehat{2}} = -0.180000$ inch
D(required) = 0.000000 inch
D'(required) = 0.000000 inch.

Upon making the indicated substitutions into Eq. [335], the results are

$M = +0.980392$
$M' = +1.020000$

which check the required values of M and M' exactly. With similar substitution into Eq. [336], the results are

$D = +0.000004$ inch
$D' = +0.000005$ inch

whereas the sixth decimal digits should be zero for both dimensions.

The checks by means of Eqs. [335] and [336] are considered highly satisfactory. Readers can use many of the equations of this paper to devise procedures for checking the solution of Example 19 in as much additional detail as desired.

VI. Some Further Questions for the Study of Example 19.

Example 19 has, no doubt, raised many questions in readers' minds, which they are encouraged to look into on their own initiative. The writer suggests they also examine the answers to the following questions.

1. In principle, the problem of Example 19 can be solved by the use of a single negative lens. Examine the results of proceeding to reach such a solution.
2. Starting with the preliminary- and primary-stage solutions that are identified by Table 19, Item 18; complete the secondary-stage solution on the assumption that $S_1 = -0.180000$ inch, that $S_2 = +0.060000$ inch, and that a, $\overline{S}$, and f_1 remain unchanged.
3. In the solution of Example 19 that is represented by Table 19, Item 18 and Table 20, Item E, change the focal length $\widehat{f}_2'$ of component lens $\widehat{L}_2'$ to $f_2' = -0.968223$ inch, and change focal length $\widehat{f}_2'$ of component lens $\widehat{L}_2'$ to $\widehat{f}_2' = -0.968223$ inch, and change focal length $\widehat{f}_2$ of component lens $\widehat{L}_2$ to $\widehat{f}_2 = +2.448515$ inches. Keeping d_2, a, and $\overline{S}_{\widehat{2}}$ unchanged, what happens to the Gaussian solution to the problem? What happens to the physical solution of the problem?

4. Assuming that nothing is changed in the specification of Example 19 except the value of $M' = 1/M$, what is the approximate limit to the value of $(M' > 0)$? Is it, for instance, greater or less than $M' = +1.1$? Greater or less than $M' = +1.5$?
5. Given Eqs. [324], [325], and [326], what are the implications of insisting that the solution of Example 19 *must* assume that $k = +1$ or that $f_2 = f_1$?

31. A Note On the Number of Decimal Places Retained In the Sample Calculations Performed In the Examples of This Paper

In applying the Gaussian theory and its equations, there is usually no practical reason to retain more than a decimal place or two. It is a concern of this paper, however, to develop exact equations that do not themselves unduly limit the accuracy of calculations that are performed with them. It is believed that this is demonstrated to have been accomplished if calculations that retain a relatively large number of decimal places are confirmed to essentially the same number of places by check calculations that employ other formulas or routines of calculation. When users are thus assured of the accuracy of the equations, they can freely be their own judge of the accuracy they will demand of their computations.

32. Brief Historical Note and Summary Conclusion

The practice and the science of European lens optics began about 1250–1300 with efforts to improve aging human vision by the use of crude spectacle lenses. This led to further experimentation with lenses, and in the late 1500s and early 1600s the powers of normal human vision were greatly extended by two inventions, the compound microscope and the telescope.

The improvements and extensions of human visual powers were real, but the images had troublesome imperfections that were inherent in the lens means that produced them. After the appearance of the microscope and telescope, image imperfections and their reduction or elimination gradually became a major object of study in the developing science of geometrical optics. This trend was reinforced by the development of the photographic camera in the early 1800s for photographic lenses exhibited the same image defects. The improvements that have since been made in the quality of images

in all lens applications have been astonishingly great.† Image quality remains a problem, however, that continues to receive extremely close attention in the theoretical and experimental development as well as design and manufacture of lenses and optical systems.

In about 1840 Carl Friedrich Gauss published his version of the theory of lens system optics, under the title "Dioptrische Untersuchungen." This book dealt with the locations and sizes of images in relation to focal lengths and object-locations in lens systems generally; the theory was restricted to object-points and image-points that were either on or very close to the lens system axis, and to rays that lie extremely (or paraxially) close to the axis. The Gaussian theory did not promise to be of much use in solving the problems of image quality, but it did supply insights that allowed the problem to be described in the broadest of terms. The theory apparently received little published attention or development after a version of it was published in Italian by Galileo Ferraris in 1876.††

Persons who become established in the practice and science of lens optics, by close association with the methods and practices of the past that are embodied in existing institutions, appear not to be limited in their accomplishments by the lack of published development of the Gaussian theory. However, persons without access to such association are not as fortunate. If they must work alone with their lens problems, they have great difficulty in comprehending and employing the simple or fundamental quantitative relationships that exist within a lens system.

This paper attempts further development of the Gaussian theory of lens systems to make it possible for the non-expert to lay out

† Sustained mostly by the importance and the wonder of the human visual sense, and of the wider worlds it might (>1609) be assisted to reveal, for nearly four centuries dedicated human effort and inspiration were devoted to solving a host of lens problems, the nature and extreme difficulty of which could certainly not have been suspected in 1300 nor, perhaps, even in 1600. The results of these centuries of effort are concentrated, for example, in a modern photographic objective, which could have had no existence without *all* of those results. Some of the most important among such results have only been attainable since the late 1930s and the early 1950s. The commanding importance of human sight in human affairs is attested by efforts to further develop and improve the means and methods that are used in conventional imaging systems (and also to develop radically unconventional systems) that continue unabated in the 1980s. This carries on the tradition begun in the late 1200s that has progressively developed without interruption since the early 1600s.

†† Originally titled "The Fundamental Principles of Dioptric Instruments." English tr. by Oscar Faber titled "Ferraris' Dioptric Instruments," London, H. M. Stationery Office (February 1919).

systems to accomplish desired imagery in a desired space and to determine focal lengths and spacings of lenses that can be used in experimental trials and tests of such systems on the optical bench.

Such further developed theory goes far beyond blindly exploratory, experimental lens-bench setups, and beyond trial calculations that are based on those narrowly restricted Gaussian relationships heretofore published which appear as Eqs. [1], [2], [32], [33], and [45] in Part 1 of this paper. The extended theory makes it possible to avoid use of the graphical and other roundabout methods, or the dependence on liberal amounts of intuition that must often supplement the incompleteness of those five equations.

The methods of this paper include the following (all references appear in Part 1 of the paper).

1. Adoption of a purely optical sign convention, that is readily manageable under all practical conditions of lens system location and spatial orientation (see Sec. 3).
2. Extension of the routinely published, two Gaussian "thin lens" equations (Eqs. [1] and [2]) to include a third equation (Eq. [3]) that effectively makes accessible the full analytical potential of the Gaussian theory (see Sec. 3).
3. The development of general equations for systems of two separated lenses (or two separated systems of lenses) that provide a basis for deriving many useful Gaussian lens system relationships (see Secs. 5, 7, and 9; Example 6; appendixes 1 and 2; etc.).
4. A simplifying replacement transformation that eliminates much unnecessary work and is a very great time and trouble saver (see Sec. 6).
5. A simplification of the analytical solutions of many lens problems by use of a theorem that has been named the Conjugate Displacement Theorem (see Sec. 10 and Example 7).
6. Complete solutions of the three basic Gaussian, "thick lens" (or system) Eqs. [1], [2], and [3] of Part 1, and their inclusion in Table 1 for easy, working reference (see Secs. 3 and 4).

A fair sampling of examples that are illustrative of practical applications of the extended Gaussian theory is included in the three parts of this paper. Frequent derivation of new equations in the text and appendices illustrates the extensive analytical possibilities of the three basic equations, as well as uses of the special methods cited in the six items above. The paper has suggested and emphasized, and here emphasizes again, that this study of the optical relationships that are implicit in the general Gaussian Eqs. [1], [2],

and [3] is by no means complete. The approach that is taken by the study offers apparently unlimited opportunity for further extension in any manner and direction that might appear promising or of interest.

To return to the question of image quality: This problem may be very broadly framed in terms of the simplest concepts of Gaussian lens theory. In actual practice the achievement of good image quality assumes and substantially requires the following:

7. All object-points are in a plane normal to the lens axis.
8. The images formed of those points are in a plane that is normal to the axis.
9. All light, regardless of its color, in the image of each point, is confined to an area that is as small as possible; a so-called "point" image. This already indicates the presence of compromise in the work of improving image quality.
10. The image of the axial object-point coincides as nearly as possible with the location of its Gaussian or paraxial image, which is formed in light of some specific, predetermined color or wavelength.
11. The magnification of the image is the Gaussian or paraxial magnification that is exhibited by the system, in item 10 above. Call it M.
12. The image of an object-point that is distant h from the axis is distant $h' = Mh$ from the axis.

Further in relation to image quality: The Gaussian theory is like a scaffold that is essential to the construction of a building, but that contributes few or none of the details. In lens system design practice, The Gaussian theory is constantly represented by paraxial ray traces, and principal ray or chief ray traces, and the interpretations that are based on them.

The theory of high-quality images is very involved. A good basic treatment of it appears in Parts One and Two of references 5 and 21 that accompany Parts 1 and 2 of this paper.

The discussion of the relation of Gaussian theory to image quality justifies the following paragraph, which is quoted from Sec. 23, Conclusion of Part 2.

"Once a system is initially conceived in Gaussian terms, its later elaboration and development may cause it to differ widely from its preliminary form in order to meet all requirements. It will usually, however, continue to conform closely in paraxial terms to the broadly stated Gaussian relationships of the initial solution."

Acknowledgment

Numerous predecessors and numerous contemporaries of the writer have all contributed in many varied ways to making it possible to write this paper. To all of them go the writer's most appreciative if wholly inadequate thanks.

Appendix 9—Outline of Procedure for Constructing Table 15 from Table 14.

In Table 14, the four columns headed "System Nos.," ". . . Conjugate Gaussian Inequalities," "*f*-signets," and "*M*-signets With Fit n.-metrics" are chiefly of interest. The footnote to Sec. 28 (*RCA Rev.* 45, No. 1, p. 136) defines and describes the terms "signet" and "metric."

Any numbered system from the above identified columns directly applies to an actual Gaussian image-forming situation, in which one object O and its image I are involved (see Fig. 16). In fact no other actual Gaussian image-forming situations exist for lenses and lens systems. Where relaying actions are involved, any of the numbered systems will similarly apply to a second object O′ and its image I′, if identifying primes are added to *X*, *Y*, and *M*. Focal length *f* remains unprimed on the assumption that it remains the same for the imagery of both O and O′.

In other words, in a relaying action O and O′ may each be imaged according to the relationships of a single system (System No. 1, for example), or, on the other hand, each may be imaged according to a different one of the two relationships that apply in Table 14 to any two differently numbered systems, *provided the two systems share the same f-signet.*

To illustrate, O may be imaged according to either System No. 2 or System No. 1 (in terms of the *X*-, *Y*-, *f*-, and *M*-signets of the system), while O′ is imaged according to System No. 5′ (in terms of the *X*′-, *Y*′-, *f*-, and *M*′-signets of that system). The *f*-signet is ($f > 0$) for all three of those systems. Note that when the object is O′, it is convenient to also prime the System No. in the interest of easy identification, and this practice will be followed. O and O′ may be arbitrarily attached to any two of those three (or other corresponding) systems, or to any one of them.

In the Conjugate Gaussian Inequalities column of Table 14, it is easy to see that the groupings or combinations of the *X*-, *Y*-, and *f*-signets are different for each of the six systems. These groupings are uniquely determined by the nn.-metrics that relate the signets of each system in this column. For example, the three indicated

signets are grouped for each system according to f-signet in the manner of Table 23 below.

The reader can easily see and take note of the fact that the two hypothetical groupings $(X > 0)$, $(Y < 0)$, $(f > 0)$ and $(X < 0)$, $(Y > 0)$, $(f < 0)$, which are absent from Tables 14 and 23, can have no existence in Gaussian imagery. For, by substituting the signets of the first of these "absent" groupings into (II-1), Table 1, Part 1, the result is

$$\frac{(X > 0)(Y < 0)}{(X > 0) - (Y < 0)} = (f > 0) \qquad [400]$$

which is an unconditionally false "equation" of sign.† Similar substitution of the second absent grouping also produces an unconditionally false "equation" of sign. On the other hand, corresponding substitutions of the signet groupings of Systems No. 1, 2, 3, and 4 result in conditionally true equations of sign, due to the corresponding denominators in Ineq. [400] each being the difference of two *like* signets in which $|X|$ may be either $\gtrless |Y|$. System Nos. 5 and 6 produce unconditionally true (identity type) equations of sign because the corresponding denominators are each the difference of two *unlike* signets. A true equation of sign merely equates expressions that have the same sign.

(In contrast to the foregoing paragraph and Table 23, the situation is quite different for the case of the column, ". . . Conjugate Newtonian Inequalities" of Table 14. Here it is found not only that the X_N, Y_N, and f-signet groupings are identical for System Nos. 1 and 2, but that the groupings are also identical for System Nos. 3 and 4. The distinctions between the two members of each of these two pairs of systems, that correspond to Table 23, probably depend

Table 23—The six unique groupings of the Gaussian X-, Y-, and f-signets of System Nos. 1 through 6 in Table 14 for the f-signets $(f > 0)$ and $(f < 0)$.

System Nos.	X-, Y-, and f-Signet Groupings
$(f > 0)$	
1	$(X < 0), (Y < 0), (f > 0)$
2	$(X > 0), (Y > 0), (f > 0)$
5	$(X < 0), (Y > 0), (f > 0)$
$(f < 0)$	
3	$(X > 0), (Y > 0), (f < 0)$
4	$(X < 0), (Y < 0), (f < 0)$
6	$(X > 0), (Y < 0), (f < 0)$

† See footnote to Ineq. [392], Appendix 8 (*RCA Rev.* **45**, No. 1, p. 163).

Table 24—Relaying-actions, of pair-combination groupings of System Nos. 1, 2, and 5 of Table 14 for the *f*-signet ($f > 0$) *only*.

Group Nos.	System No. Combinations (See Table 14)	Correlations of X-, Y-, X'-, Y'-, M-, and M'-Signets, Together with Applicable M- and M'-n.-Metrics, When the f-Signet is ($f > 0$)		
I_a	No. 1 with No. 1′.	$(X < 0),(Y < 0)$	$(X' < 0),(Y' < 0)$	$(M > 0) > 1$ $(M' > 0) > 1$
I_b	No. 1 with No. 2′.	$(X < 0),(Y < 0)$	$(X' > 0),(Y' > 0)$	$(M > 0) > 1$ $(M' > 0) < 1$
I_c	No. 2 with No. 1′.	$(X > 0),(Y > 0)$	$(X' < 0),(Y' < 0)$	$(M > 0) < 1$ $(M' > 0) > 1$
I_d	No. 2, with No. 2′.	$(X > 0),(Y > 0)$	$(X' > 0),(Y' > 0)$	$(M > 0) < 1$ $(M' > 0) < 1$
II_a	No. 1 with No. 5′.	$(X < 0),(Y < 0)$	$(X' < 0),(Y' > 0)$	$(M > 0) > 1$ $(M' < 0)$
II_b	No. 2, with No. 5′.	$(X > 0),(Y > 0)$	$(X' < 0),(Y' > 0)$	$(M > 0) < 1$ $(M' < 0)$
III_a	No. 5 with No. 1′.	$(X < 0),(Y > 0)$	$(X' < 0),(Y' < 0)$	$(M < 0)$ $(M' > 0) > 1$
III_b	No. 5 with No. 2′.	$(X < 0),(Y > 0)$	$(X' > 0),(Y' > 0)$	$(M < 0)$ $(M' > 0) < 1$
IV	No. 5 with No. 5′.	$(X < 0),(Y > 0)$	$(X' < 0),(Y' > 0)$	$(M < 0)$ $(M' < 0)$

on the simple nn.-metrical relationships between the signets, and may depend also on n.-metrical relationships that this paper does not define and examine. A result of this situation is that the Newtonian lens equations identifying fewer distinctions of optical imagery that depend *only on the signs* of the quantities involved than do the Gaussian lens equations, which makes the Newtonian equations less generally useful.)

Table 24 presents the nine System No. combinations that may be involved in a relaying action, given the signet groupings of the X-, Y-, X'-, Y'-, M-, and M'-signets that are represented by the "Gaussian nn.-metrics; . . ." of Table 14, *when the f-signet is* ($f > 0$). Each combination is given a Roman group-identification number with a lower case literal subscript (there is no subscript for the unique Group IV). The Roman group-numbers are assigned in Table 24, to correlate the table with Table 12 according to the occurrence in them of the magnification-signets, as noted in the Table 25.

The scheme of Table 24 may easily be extended by the reader for those system combinations where the *f*-signet is ($f < 0$). The groups of such an extension will comprise, from Table 14:

Table 25—The M-, M'-, and M_Δ-signets that correlate Groups I_a through IV of Table 24 (for the signet ($f > 0$) only), with the same Group Numbers of Table 12.

In Groups I, $(M > 0),(M' > 0),(M_\Delta > 0) = MM'$; as in Group I, Table 12.
In Groups II, $(M > 0),(M' < 0),(M_\Delta < 0) = MM'$; as in Group II, Table 12.
In Groups III, $(M < 0),(M' > 0),(M_\Delta < 0) = MM'$; as in Group III, Table 12.
In Group IV, $(M < 0),(M' < 0),(M_\Delta > 0) = MM'$; as in Group IV, Table 12.

Firstly, Group No. V, System Nos. 6 and 6′.†

Secondly, Group No. VI_a, System Nos. 6 and 4′ and Group No. VI_b, System Nos. 6 and 3′.

Thirdly, Group No. VII_a, System Nos. 4 and 6′, and Group No. VII_b, System Nos. 3 and 6′.

Fourthly, Group No. $VIII_a$, System Nos. 4 and 4′, and Group No. $VIII_b$, System Nos. 4 and 3′, and also Group No. $VIII_c$, System Nos. 3 and 4′, and Group No. $VIII_d$, System Nos. 3 and 3′.

In such an extension of Table 24 the correlation between that table and Table 12, according to their Group Nos. V, VI, VII, and VIII, parallels the general scheme of Table 25.

The information collected in Table 24 makes it possible to determine the signs of the displacements $\Delta_c x$ and $\Delta_c y$ for every system combination for which the f-signet is ($f > 0$). The general procedure for identifying the $\Delta_c x$- and $\Delta_c y$-signets is outlined in summary form in Table 26. In this table $\Delta_c x$ and $\Delta_c y$ are expressed in terms of (X,X') and (Y,Y'), respectively, as they are in Eqs. [172] and [173]. It is immaterial whether x and x' or X and X' are used in Eq. [172] and correspondingly for Eq. [173] (see Fig. 10 and Sec. 17, Part 2).

In Table 26, Group No. I_a, the signs of $\Delta_c x$ and $\Delta_c y$ are both inherently indeterminate because they depend only on the differences between negative signets whose relative numerical values are unknown.†† However, the signs of $\Delta_c x$ and $\Delta_c y$ *must* be the same because $\Delta_c y/\Delta_c x = MM' = (M_\Delta > 0)$, (see Table 14). With this restriction, the signs of $\Delta_c x$ and $\Delta_c y$ in Group No. I_a are arbitrary in practice. Similar considerations also apply in the cases of Group Nos. I_d and IV. In no other groups of Table 26 are the $\Delta_c x$- and $\Delta_c y$-signets both inherently indeterminate. Extension of Table 26 to

† To avoid confusion, readers *must* note that the Roman numeral "Table 13 Case Ident. Nos." in the second column of Table 14, are *not* related to the Roman numeral Group Nos. of Table 12 and Tables 24, 25, 26, and 15.

†† All differences between like signets, whether both positive or both negative, are indeterminate in sign. Any difference between unlike signets is always determinate in sign and expressible as a signet.

Table 26—The M_Δ-signets with the $\Delta_c x$- and $\Delta_c y$-signets and their status for Groups No. I_a through IV; where the f-signet is ($f > 0$), *only;* reference Tables 25 and 14.

Group Nos.	System No. Combinations (See Table 24)	M_Δ-Signets ($M_\Delta = MM' = \Delta_c y/\Delta_c x$)	$\Delta_c x$- & $\Delta_c y$-Signets and their Status, (See Eqs. [172] and [173], Part 2)
I_a	No. 1 with No. 1′	$(M_\Delta > 0)$ By Table 14 $< \begin{matrix}(M > 0)\\(M' > 0)\end{matrix}$	$\Delta_c x = (X' < 0) - (X < 0) \gtrless 0$, indeterminate. $\Delta_c y = (Y' < 0) - (Y < 0) \gtrless 0$, because $(M_\Delta > 0)$.
I_b	No. 1 with No. 2′	$(M_\Delta > 0)$ By Table 14 $< \begin{matrix}(M > 0)\\(M' > 0)\end{matrix}$	$\Delta_c x = (X' > 0) - (X < 0) > 0$, determinate. $\Delta_c y = (Y' > 0) - (Y < 0) > 0$, determinate.
I_c	No. 2 with No. 1′	$(M_\Delta > 0)$ By Table 14 $< \begin{matrix}(M > 0)\\(M' > 0)\end{matrix}$	$\Delta_c x = (X' < 0) - (X > 0) < 0$, determinate. $\Delta_c y = (Y' < 0) - (Y > 0) < 0$, determinate.
I_d	No. 2 with No. 2′	$(M_\Delta > 0)$ By Table 14 $< \begin{matrix}(M > 0)\\(M' > 0)\end{matrix}$	$\Delta_c x = (X' > 0) - (X > 0) \gtrless 0$, indeterminate. $\Delta_c y = (Y' > 0) - (Y > 0) \gtrless 0$, because $(M_\Delta > 0)$.
II_a	No. 1 with No. 5′	$(M_\Delta < 0)$ By Table 14 $< \begin{matrix}(M > 0)\\(M' < 0)\end{matrix}$	$\Delta_c y = (Y' > 0) - (Y < 0) > 0$, determinate. $\Delta_c x = (X' < 0) - (X < 0) < 0$, because $(M_\Delta < 0)$.
II_b	No. 2 with No. 5′	$(M_\Delta < 0)$ By Table 14 $< \begin{matrix}(M > 0)\\(M' < 0)\end{matrix}$	$\Delta_c x = (X' < 0) - (X > 0) < 0$, determinate. $\Delta_c y = (Y' > 0) - (Y > 0) > 0$, because $(M_\Delta < 0)$.
III_a	No. 5 with No. 1′	$(M_\Delta < 0)$ By Table 14 $< \begin{matrix}(M < 0)\\(M' > 0)\end{matrix}$	$\Delta_c y = (Y' < 0) - (Y > 0) < 0$, determinate. $\Delta_c x = (X' < 0) - (X < 0) > 0$, because $(M_\Delta < 0)$.
III_b	No. 5 with No. 2′	$(M_\Delta < 0)$ By Table 14 $< \begin{matrix}(M < 0)\\(M' > 0)\end{matrix}$	$\Delta_c x = (X' > 0) - (X < 0) > 0$, determinate. $\Delta_c y = (Y' > 0) - (Y > 0) < 0$, because $(M_\Delta < 0)$.
IV	No. 5 with No. 5′	$(M_\Delta > 0)$ By Table 14 $< \begin{matrix}(M < 0)\\(M' < 0)\end{matrix}$	$\Delta_c x = (X' < 0) - (X < 0) \gtrless 0$, indeterminate. $\Delta_c y = (Y' > 0) - (Y > 0) \gtrless 0$, because $(M_\Delta > 0)$.

include the groups for which the f-signet is ($f < 0$) will produce corresponding inherent indeterminacies. Discussions that relate to such an extension of the table will appear below. It develops that in *all* groups where the $\Delta_c x$- and $\Delta_c y$-signets are *both* inherently indeterminate, they *must both* be given the *same* arbitrary sign in practice.

In Group Nos. II_a, II_b, III_a, and III_b of Table 26, the sign of either the $\Delta_c x$- or the $\Delta_c y$-signet is inherently *determinate* while that of its conjugate displacement-signet, although inherently *indeterminate,* is functionally assignable on the basis of the conjugate displacement theorem of Eq. [232]. Thus, when either signet is determinate, the M- and M'-signets of Table 25 (or Table 14) and the relationship of Eq. [232] make it easy to assign or identify the conjugate signet.

As in the case of Table 24, Table 26 includes only Group Nos. I_a through IV for which ($f > 0$). The scheme of Table 26 may be carried forward by the reader in the cases of all System No. combinations for which the f-signet is ($f < 0$), after Tables 24 and 25 have been extended in the manner suggested above.

Table 15, whose construction depends on the procedures outlined in this appendix and in Appendix 8, is largely assembled from the information that is collected in Tables 24 and 26, for Groups I_a through IV. The reader's extensions of those tables will provide the information for Group Nos. V through $VIII_d$, from which the corresponding portions of Table 15 may be similarly assembled.

For example, in Table 15, the line for Group No. I_d, System Nos. 2 with 2′, pertains to ($f > 0$) for column (1), and is completed for columns (2)* through (7)′ from the information assembled for Group No. I_d in Table 24. The same line, in Table 15, is similarly completed for columns (8), (10), and (11) from the information assembled for Group No. I_d in Table 26.

Signets do not appear as such in Table 15. For example, in column (4)* the letter P is a convenient contraction of the positive signet ($Y > 0$) to indicate that Y is positive, and in column (11) the letter N is a convenient contraction of the negative signet ($\Delta_c y < 0$) to indicate that $\Delta_c y$ is negative. This practice is followed generally throughout Table 15.

In Table 15, when a signet is arbitrary, the fact is indicated as in columns (10) and (11) by the contraction P/N; but the P in either of these two columns *always* implies P in the other, and similarly for N.

In columns (6)* and (7)′, n.-metrics are used to indicate further limitations on all of the *positive* M- and M'-signets; in Group No.

I_d these *must* both be less than unity, as symbolized by the $P < 1$ entries in those columns of Table 15. (The n.-metric $P = 1$ is a special case not included in columns (6)* and (7)′ of Table 15, but it is always easily understood to signify that $X = Y = 0$ or that $X' = Y' = 0$, as demonstrated in Sec. 8, Part 1. Inclusion of these simple cases would have greatly complicated Table 15). No special limitations apply to the *negative* M- and M'-signets of columns (6)* and (7)′.

In column (8), the P and N values assigned to M_λ correspond, respectively, to the "regular" and "inverse" relaying actions of column (9), as defined in Sec. 27.

Note is taken here that in Table 15 the distinction between the two members of each of the following sets of Group Nos. $\{I_b,I_c\}$, $\{II_a,III_a\}$, $\{II_b,III_b\}$, $\{VI_a,VII_a\}$, $\{VI_b,VII_b\}$, and $\{VII_b,VIII_c\}$ is only nominal and practical and is not fundamental as is, for example, that of the set $\{I_b,III_a\}$. The same is noted of the distinction in Table 12, between Group Nos. II and III, and between Group Nos. VI and VII, which in both cases is nominal and practical, and not fundamental.

Appendix 10—Occurrence and Resolution of Ambiguous Solutions in the Practical Measurement of the Index of Refraction of a Single Thick Lens by the Methods of Sec. 30A, Example 18.

The conjugate solutions of Eq. [287] that are expressed by Eq. [309] may be more compactly and, for some purposes, more usefully expressed as

$$n_1 = A_n + (A_n^2 - B_n)^{1/2} \qquad [401]$$

and

$$n_2 = A_n - (A_n^2 - B_n)^{1/2} \qquad [402]$$

where, as in Eq. [309], the *fixed* values of A_n and B_n in the case of any actual lens, are

$$A_n = \frac{\overline{F}(r_2 - r_1 + 2d) + r_1 r_2}{2\overline{F}(r_2 - r_1 + d)} \qquad [403]$$

and

$$B_n = \frac{d}{r_2 - r_1 + d}. \qquad [404]$$

Also, the useful relationships

$$\frac{n_1 + n_2}{2} = A_n \qquad [405]$$

and

$$n_1 n_2 = B_n \qquad [406]$$

are easily derived from Eqs. [401] and [402].

The numerical subscripts added to the n's merely indicate each of them to be a different one of the two real conjugate solutions of Eq. [287], which is quadratic in n.

The quantities that determine A_n and B_n are the three fixed geometric dimensions of the lens and, in the case of A_n, the fixed focal length of the lens which, in turn, depends on the identity of the glass and the wavelength of the light. Furthermore, the measurements of the quantities M, M', $\Delta_c x$ ($= x' - x$, both arbitrary), or $\Delta_c y$ ($= y' - y$, both arbitrary) that express the focal length in Eq. [310] do not in any way affect the calculated value of focal length $\overline{F}$, except as a result of measurement inaccuracies. In other words, the calculated index values n_1 and n_2 that are expressed by Eqs. [401] and [402] are *both* fixed and independent of the measurements except for measurement errors. As a matter of practical information, the refractive index of *any* optical glass (in the present state of the art) typically lies somewhere inside an interval that extends roughly between $+1.4$ and $+2.0$.

The remainder of this appendix deals with the case of a lens whose actual glass index is $n_2 = 1.5407$, but whose geometry and focal length are such that the second value of index, calculated by Eq. [401], is $n_1 = 1.6236$. In a case of this kind the two refractive indexes that are yielded by the measurements of "Possible Procedure No. 1," and represented by the development of Eq. [309] and its dependent Eqs. [401] through [406] *both* lie inside the refractive index interval that also includes *all* existing optical glasses. It is clearly a shortcoming of Procedure No. 1 that the required measurements are incapable of identifying which of n_1 and n_2 is the true index of the lens. Since it must be the one or the other, the procedure is inherently capable of producing an ambiguous result that it cannot resolve.

A lens that will lead to this difficulty has the following specification:

$r_1 = +0.3043640345$ u (where u is an arbitrary unit)
$r_2 = -0.148135228$ u
$d = +0.750 \ldots$ u
$\overline{F} = +0.4390 \ldots$ u

The focal length $\bar{F}$ of this lens is unaffected by changing the glass index from $n_2 = 1.5407$ to $n_1 = 1.6236$, but the value of $\bar{F}$ is different for all other values of n. These statements are easily checked by making substitutions into Eq. [287]. The specification for this lens and the procedure for arriving at it are clarified and developed in the final paragraphs of this appendix that are associated with Eqs. [421] through [434].

"Possible Procedure No. 2" of Example 18 is entirely different in its origin and form from Procedure No. 1. It derives from Eqs. [19] and [312], where any arbitrary and directly measurable vertex object distance x_1 of any simple thick lens is related to its conjugate and directly measurable vertex image distance $\bar{y}_2$ by a function that involves the two constant individual focal lengths f_1 and f_2 of its plano-components L_1 and L_2 and their constant optical separation $\bar{a}$ (see Sec. 30). Each of f_1, f_2, and $\bar{a}$ is restated in terms of r_1 and k, r_2 and k, and d and k, respectively, to produce Eq. [313]. This equation is quadratic in k, where $k = 1/(n - 1)$. Thus, k represents the constant refractive index of the lens in, say sodium-d light, but is not equal to the index. The equation relates $\bar{x}_1$ to $\bar{y}_2$ by means of the constant quantities r_1, r_2, d, and k that represent the geometry of the lens and its material. When reduced to standard form, the quadratic in k becomes Eq. [315], whose conjugate solutions are represented by Eqs. [316] and [317]. In Procedure No. 2, the focal length of the lens under measurement and the magnification at which any set of measurements is made do not become explicitly involved.

The solutions that are expressed by Eq. [316] may be more compactly and for some purposes more usefully expressed as

$$k_1 = A_k \pm (A_k^2 - B_k)^{1/2}, \text{ or more generally as}$$
$$k_{\pm V} = A_k \pm (A_k^2 - B_k)^{1/2} \qquad [407]$$

and furthermore as

$$k_2 = A_k \mp (A_k^2 - B_k)^{1/2}, \text{ or more generally as}$$
$$k_{\mp V} = A_k \mp (A_k^2 - B_k)^{1/2} \qquad [408]$$

where, as in Eq. [316], the values of A_k and B_k are

$$A_k = \frac{x_1 y_2(r_1 - r_2) + r_1 r_2(x_1 - y_2) - (x_1 r_2 + y_2 r_1)d}{-2r_1 r_2(x_1 - y_2 - d)} \qquad [409]$$

and

$$B_k = \frac{\bar{x}_1 \bar{y}_2(r_1 - r_2 - d)}{r_1 r_2(x_1 - \bar{y}_2 - d)}. \qquad [410]$$

Also, following Eq. [317],

$$n_1 = \frac{1}{k_{\pm\sqrt{}}} + 1 = \frac{1 + k_{\pm\sqrt{}}}{k_{\pm\sqrt{}}} \qquad [411]$$

and

$$n_2 = \frac{1}{k_{\mp\sqrt{}}} + 1 = \frac{1 + k_{\mp\sqrt{}}}{k_{\mp\sqrt{}}} \qquad [412]$$

in which n_1 and n_2 represent cases of reversed signs before the radical in Eq. [407] *and* in Eq. [408] and have no other significance.

In terms of any actual lens that is measured, the quantities r_1, r_2, d, and k of Eqs. [313], [409], and [410] are all constant, and in Eq. [313] the measured vertex conjugate distance $\bar{y}_2$ is a rational function of the *single*, independent or arbitrary, measured variable $\bar{x}_1$. It follows that for every arbitrary, measured value of $\bar{x}_1$, and of its measured dependent conjugate $\bar{y}_2$, *one* of the two solutions of quadratic Eq. [313] for k *must* be a constant function† (except for measurement errors, whose effect is not discussed here) of the single independent variable $\bar{x}_1$. Let this solution be represented by, say $k_{\pm\sqrt{}}$ of Eq. [407], which assumes that no errors exist in the measurements.

The second or conjugate solution of Eq. [313] for k is then represented by $k_{\mp\sqrt{}}$ of Eq. [408]. Should this solution also be a constant function of $\bar{x}_1$, the potential ambiguity of Procedure No. 1 would also reside in Procedure No. 2. It is left to the reader to show mathematically that when the solution of Eq. [407] is a constant function of $\bar{x}_1$, the solution of Eq. [408] is a variable (or at least not necessarily a constant) function of $\bar{x}_1$. The reader will find it relatively easy to demonstrate that if it is assumed that A_k and B_k is each a variable function of $\bar{x}_1$ (as is readily conceivable), then $k_{\mp\sqrt{}}$ *must* be a variable function of $\bar{x}_1$ if $k_{\pm\sqrt{}}$ is a constant function of it. This may be helpful in seeking the more general proof.

For the less theoretical purposes of this appendix, the properties of Procedure No. 2 will now be developed by making precise numerical substitutions into Eqs. [407] through [412]. These substitutions show empirically that $k_{\mp\sqrt{}}$ *may* be a variable function of $\bar{x}_1$

† See R. Courant and H. Robbins, *What is Mathematics?* p. 275 (Oxford University Press, NY, 4th Ed., 1947–51). The interesting and useful topic of a constant function of a variable parameter appears to be traditionally ignored by textbooks that develop the calculus. Such constant functions tend to be loosely regarded as "independent" of the parameter and, therefore, unimportant from the elementary calculus point of view. This misapprehension needs to be corrected by the "expert" writers of calculus texts.

when $k_{\pm\vee}$ is a constant function of it. When this is true, no undetectable ambiguity will occur. The effects of errors that are made in measuring the five quantities in Eqs. [409] and [410] are not considered.

The lens will be the same that was specified above to have $\overline{F} = +0.4390\ldots$ u whether the glass index is $n_1 = 1.6236$ or $n_2 = 1.5407$. The glass selected for the lens has index $n_2 = 1.5407$. Values of set $\{\widehat{x}_1,\widehat{y}_2\}$ are calculated to correspond to an arbitrary $M = -2.0\ldots$, and of set $\{\widehat{x}_1',\widehat{y}_2'\}$ to correspond to an arbitrary $M' = -15.0\ldots$. They will be precisely calculated for the selected n_2 and for the lens dimensions

$r_1 = +0.3043640345$ u (where u is an arbitrary unit)
$r_2 = -0.1458135228$ u
$d = +0.750\ldots$ u

By Eqs. [279], [280], and [282]

$$\widehat{f}_1 = r_1/(n_2 - 1) = +0.5629074061 \text{ u}, \qquad [413]$$

$$\widehat{f}_2 = -r_2/(n_2 - 1) = +0.2696754629 \text{ u}, \qquad [414]$$

$$\widehat{a} = d/n_2 = +0.4867917181 \text{ u}. \qquad [415]$$

By solution of Eq. [13] of Part 1

$$\widehat{x}_1 = \frac{\widehat{f}_1\widehat{f}_2 - M\widehat{f}_1(\widehat{f}_2 - \widehat{a})}{M(\widehat{f}_1 + \widehat{f}_2 - \widehat{a})} \qquad [416]$$

and by replacement transformation of this solution, according to Table 3, Sec. 6, Part 1

$$\widehat{y}_2 = \frac{-M\widehat{f}_1\widehat{f}_2 + \widehat{f}_2(\widehat{f}_1 - \widehat{a})}{(\widehat{f}_1 + \widehat{f}_2 - \widehat{a})}. \qquad [417]$$

Thus, for arbitrary $M = M = -2.0\ldots$; $\widehat{x}_1 = +0.1339397791$ u, $\widehat{y}_2 = +0.9373610723$ u and for arbitrary $M = M' = -15.0\ldots$; $\widehat{x}_1' = +0.3241731127$ u, $\widehat{y}_2' = +6.644361075$ u

Equation [40], Part 1, may be used as a check, thus

$$\frac{\widehat{y}_2' - \widehat{y}_2}{\widehat{x}_1' - \widehat{x}_1} = MM' = \frac{6.644361075 - 0.9373610723}{0.3241731127 - 0.1339397791} = +29.99999997 \qquad [418]$$

which is considered an excellent check.

The values of $k_{\pm\vee}$, $k_{\mp\vee}$, n_1, and n_2 are calculated from Eqs. [407] through [412]. Those of $k_{\pm\vee}$ and $k_{\mp\vee}$ are checked by the relationships

$$\frac{k_{\pm\vee} + k_{\mp\vee}}{2} = A_k \qquad [419]$$

and

$$k_{\pm\vee}k_{\mp\vee} = B_k \qquad [420]$$

which are very easily derived from Eqs. [407] and [408].

By using appropriate signs before the radical and substituting the foregoing values of r_1, r_2, d, $\bar{x}_1$ and $\bar{y}_2$ into Eqs. [407] through [412], the results may be represented as (for arbitrary $M = -2.0\ldots$)

$$\left.\begin{array}{l} k_{\pm\vee} = +1.849454410,\ n_2 = +1.540700000 \\ k_{\mp\vee} = -0.2952273660,\ n_1 = -2.387219869 \end{array}\right\|$$

$$A_k = +0.7771135220;\ \text{ck.}\ \frac{k_{\pm\vee} + k_{\mp\vee}}{2} = +0.7771135220$$

$$B_k = -0.5460095537;\ \text{ck.}\ k_{\pm\vee}k_{\mp\vee} = -0.5460095540$$

and by Eq. [287], $\bar{F} = +2.133941980 \times 10^{-2}$ u when $n = n_1$; $\bar{F} = +0.4390000002$ u when $n = n_2$. Then, by changing $\bar{x}_1$ to $\bar{x}_1'$ and $\bar{y}_2$ to $\bar{y}_2'$, the results may be represented as (for arbitrary $M' = -15.0\ldots$)

$$\left.\begin{array}{l} k_{\pm\vee} = +1.849454408,\ n_2 = +1.540700001 \\ k_{\mp\vee} = -1.112829502,\ n_1 = +.1013897473 \end{array}\right\|$$

$$A_k' = +0.3683124531;\ \text{ck.}\ \frac{k_{\pm\vee} + k_{\mp\vee}}{2} = +0.3683124530$$

$$B_k' = -2.058127427;\ \text{ck.}\ k_{\pm\vee}k_{\mp\vee} = -2.058127428$$

and by Eq. [287], $\bar{F} = -6.958602049 \times 10^{-3}$ u when $n = n_1$; $\bar{F} = +0.4390000002$ u when $n = n_2$.

Thus the calculated value of $n_2 = 1.5407$ is shown empirically to be a substantially constant function of $\bar{x}_1$ while the calculated value of n_1 is seen to be a variable function of $\bar{x}_1$.

The procedure for determining r_1, r_2, d, and $\bar{F}$ for a simple thick lens so that the arbitrary refractive indexes n_1 and n_2 will *each* satisfy Eq. [287] will now be briefly indicated. By noting Eq. [404] and solving Eq. [406] for r_2

$$r_2 = \frac{d + n_1n_2(r_1 - d)}{n_1n_2}. \qquad [421]$$

Then, by substituting this value of r_2 into Eq. [405] (after noting Eq. [403]), the resulting equation readily reduces to the standard form of a quadratic in r_1

$$n_1 n_2 r_1^2 - d(n_1 n_2 - 1) r_1 - \bar{F} d(n_1 + n_2 - n_1 n_2 - 1) = 0 \qquad [422]$$

whose solution may be expressed as

$$r_{1c} = r_1 = \frac{d(n_1 n_2 - 1)}{2 n_1 n_2} \pm \left[\left[\frac{d(n_1 n_2 - 1)}{2 n_1 n_2} \right]^2 - \frac{\bar{F} d(1 + n_1 n_2 - n_1 - n_2)}{n_1 n_2} \right]^{1/2}. \qquad [423]$$

This may in turn be re-expressed in the more compact and, for some purposes, more useful forms

$$r_{1c} = r_1 = A_r \pm (A_r^2 - B_r)^{1/2} \qquad [424]$$

and

$$r'_{1c} = A_r \mp (A_r^2 - B_r)^{1/2} \qquad [425]$$

where, on noting Eq. [423]

$$A_r = \frac{d(n_1 n_2 - 1)}{2 n_1 n_2} \qquad [426]$$

and

$$B_r = \frac{\bar{F} d(1 + n_1 n_2 - n_1 - n_2)}{n_1 n_2}. \qquad [427]$$

Once again, as in the case of Eqs. [401] and [402] (also the case of Eqs. [407] and [408]), the relationships

$$(r_{1c} + r'_{1c})/2 = A_r \qquad [428]$$

and

$$r_{1c} r'_{1c} = B_r \qquad [429]$$

are easily derived from Eqs. [424] and [425].

Since merely interchanging the signs of the irrational term between Eqs. [424] and [425] interchanges the algebraic values of r_{1c} and r'_{1c}, it follows that if r_1 of the lens $= r_{1c}$ then $r_2 = -r'_{1c}$. This being the case, Eqs. [428] and [429] may be re-expressed in the form

$$(r_{1c} + r'_{1c})/2 = (r_1 - r_2)/2 = A_r \qquad [430]$$

and

$$r_{1c} r'_{1c} = -r_1 r_2 = B_r. \qquad [431]$$

Recalling, once again, that by arbitrary choice $n_1 = 1.6236$ and $n_2 = 1.5407$, then Eq. [424] becomes (by noting Eqs. [426] and [427])

$$r_{1c} = r_1 = [0.3001183715d \pm (9.007103691 \times 10^{-2}d^2 - 0.1347923829\bar{F}d)^{1/2}]\text{u} \quad [432]$$

in which d and $\bar{F}$ may be given arbitrary values (in the same unit, of course), restricted only by the requirements that $d > 0$ and that the radicand in Eq. [432] be $\geqslant 0$ so that r_{1c} will be real. In general, r_{1c} will be real if, in Eq. [423],

$$\bar{F} \leqslant \left[\frac{d(n_1n_2 - 1)}{2n_1n_2}\right]^2 \cdot \frac{n_1n_2}{d(1 + n_1n_2 - n_1 - n_2)}.\dagger \quad [433]$$

This means that if the arbitrary value of d is $d = +0.750 \ldots$ u, then $\bar{F} \leqslant +0.5011653938$ u. With this as a guide, the value of $\bar{F}$ has been arbitrarily taken as $\bar{F} = +0.4390 \ldots$ u. Equation [432] then becomes simply

$$r_{1c} = r_1 = +0.2250887786 \text{ u} \pm (5.066495825 \times 10^{-2} - 4.438039206 \times 10^{-2})^{1/2} \text{ u} \quad [434]$$

From Eq. [434], it follows that

$$r_{1c} = r_1 = +0.3043640345 \text{ u} \;\|\; -r'_{1c} = r_2 = -0.1458135228 \text{ u}$$

$A_r = +0.2250887786$; ck. $(r_{1c} + r'_{1c})/2 = +0.2250887787$

$B_r = +4.438039206 \times 10^{-2}$; ck. $r_{1c}r'_{1c} = +4.438039208 \times 10^{-2}$.

It is seen that Eqs. [430] and [431] serve very well as check formulas.

To summarize—

$n_1 = +1.6236$ Assumed arbitrarily
$n_2 = +1.5407$ Assumed arbitrarily
$d = +0.750 \ldots$ u A reasonable assumption (u is an arbitrary unit)
$\bar{F} = +0.4390 \ldots$ u A reasonable assumption (see Eq. [433] et seq.]
$r_1 = r_{1c} = +0.3043640345$ u Calculated by Eq. [424]
$r_2 = -r'_{1c} = -0.1458135228$ u Calculated by Eq. [425]

The earlier part of this appendix, associated with Eqs. [401] through [420], depended upon these assumed and calculated simple lens specifications to compare the properties of Possible Procedures No. 1 and No. 2 of Example 18.

† The $\leqslant$ sign in Eq. [433] is fixed because the second factor on the right is always positive.

Appendix 11—Outline of a Further Extension of the Theory of a Simple, Thick, Spherical Lens, Based On the Assumption That the Lens Is a System of Two Quasi-Separated Lenses In the Manner of Section 30.

This appendix suggests how a new and useful equation respecting a thick simple lens may be developed. Details are left to the reader. From Eq. [172] Part 2, $x' = x + \Delta_c x$. On substitution of this value of x' into Eq. [180], allowing $x = x_1$, $y = y_2$,

$$a = \frac{Mx_1(y_2 + MM'\Delta_c x) - M'y_2(x_1 + \Delta_c x)}{MM'\Delta_c x}. \qquad [435]$$

This equation relates some of the literal quantities that appear in the system of Fig. 12.

By assuming $M = +1 = 1/M$, the relationships (III-8) and (IV-8) of Table 1, Part 1 show that the Gaussian object- and image-distances are, respectively, $X = 0$ and $Y = 0$, or $x_1 = x_{1p}$ and $y_2 = y_{2p}$ in Fig. 12.

By also assuming $M' = -1 = 1/M'$, (III-8) of Table 1 shows that the corresponding Gaussian object-distance is $X' = -2\overline{F}$, where $\overline{F}$ is the focal length of the system of Fig. 12. This makes $\Delta_c x = x_1' - x_1 = X' - X = -2\overline{F} - 0 = -2\overline{F}$ in Fig. 12. Substitution of $M = +1$, $M' = -1$, $x_1 = x_{1p}$, $y_2 = y_{2p}$, and $\Delta_c x = -2\overline{F}$ into Eq. [435] produces

$$a = \frac{x_{1p}(y_{2p} + 2\overline{F}) + y_{2p}(x_{1p} - 2\overline{F})}{2\overline{F}}. \qquad [436]$$

Since $M = +1 = 1/M$ and $M' = -1 = 1/M'$, these special values of M and M' that produce Eq. [436] from Eq. [435] preserve the generality in Eq. [436] that makes it subject to the replacement transformation of Sec. 6, Part 1. (See derivation of Eqs. [32] and [33] from Eq. [31] Part 1.)

By further inspection of Fig. 12 it is easily seen that

$$y_{2p} = x_{1p} + \overline{S} - s_1 - s_2 - a. \qquad [437]$$

In order to now treat the system of Fig. 12 as being two quasi-separated lenses in the manner of Sec. 30, the following substitutions are made into Eqs. [436] and [437]. (See Fig. 18 and Eqs. [281] and [282].)

$a = \widehat{a} = d/n \ (d \geq 0)$ $\qquad$ $s_1 = \widehat{s}_1$
$x_{1p} = \widehat{x}_{1p}$ $\qquad$ $s_2 = \widehat{s}_2$
$y_{2p} = \widehat{y}_{2p}$ $\qquad$ $\widehat{s}_1 + \widehat{s}_2 + \widehat{a} = d$
$\overline{S} = \overline{S}$

By introducing these substitutions and equivalents, first into Eq. [437] and then into Eq. [436], the center thickness d of a simple lens is easily solved for as

$$d = n\left[\frac{\hat{x}_{1p}^2 + \overline{S}(\hat{x}_{1p} - \overline{F})}{\overline{F} + n(\hat{x}_{1p} - \overline{F})}\right] \quad (d \geq 0). \qquad [438]$$

In Eq. [438] the replacement transformation calls only for $\hat{x}_{1p}$ to be replaced by $-\hat{y}_{2p}$. After such replacement the equation still holds true. The same holds true for Eqs. [439] and [440] below, *provided* $\hat{f}_1$ and $\hat{f}_2$ are also interchanged between those two equations.

Equations [438] can be solved for $\hat{x}_{1p}$ or for any of its other variables. However, if $d = 0$ the lens is "thin," and in that case $\hat{a} = 0$ and the solution for $\hat{x}_{1p}$ *must* be $\hat{x}_{1p} = 0$ (and also $\overline{S} = 0$), and *not* the *general* quadratic solution for x_{1p} when the numerator of Eq. [438] is zero (see Fig. 18, Eqs. [437], [281], [282], [290a], and Eq. [32] Part 1).

The following equations are adapted from Eqs. [65] and [66] of Example 6, Part 1:

$$f_1 = \hat{f}_1 = (d/n)(\hat{x}_{1p} - \overline{F})/(\hat{x}_{1p} - d/n) \quad (d > 0; f_1 \neq 0) \qquad [439]$$

$$f_2 = \hat{f}_2 = (d/n)\overline{F}/\hat{x}_{1p} \quad (d > 0; f_2 \neq 0). \qquad [440]$$

Furthermore, by Eqs. [279] and [280],

$$r_1 = (n - 1)\hat{f}_1 \qquad [441]$$

$$r_2 = -(n - 1)\hat{f}_2. \qquad [442]$$

Once again, see the statement following Eq. [438] above regarding replacement transformations of Eqs. [438], [439], and [440].

The numerical values in Table 20 are easily verified by making appropriate substitutions into Eqs. [438] through [442]. In those equations, however, $\hat{f}_1$ and $\hat{f}_2$ are represented, respectively, by $\hat{f}_2$ and $\hat{f}_2'$ of Table 20 and r_1 and r_2 of the equations are represented, respectively, by r_2 and r_2' of the Table.

Equation [438] makes it easily possible to calculate the center thickness d that any simple lens must have, assuming that the refractive index n, the principal point locations as represented by $\hat{x}_{1p}$ and $\overline{S}$, and the lens' paraxial focal length $\overline{F}$ are all given or known parameters. The same will be true in terms of $\hat{y}_{2p}$ if $\hat{x}_{1p}$ is replaced by $-\hat{y}_{2p}$ in the equation. If the calculated value of d is negative, the lens is, paraxially speaking, an absolute physical impossibility unless its constants are changed. Care *must always* be taken to correctly apply the Gardner sign convention.

Equation [438] may easily be made the basis for a graphical study of the basic parameters of simple thick lenses.

Equation [438] is the result of analyzing and solving the problem of the thickness of a simple lens on its lowest and simplest terms, with all extraneous considerations eliminated (see, for example, Eq. [295]). The derivation procedure for Eq. [438] and the brief items of discussion that occur in this appendix reflect the broad unity of treatment of the simple Gaussian lens equations that exists throughout all parts of this paper.

Appendix 11 should not be terminated without noting that by substituting the value of y_{2p} from Eq. [437] into Eq. [436], the solution for a, in a system of two separated lenses in air or vacuum, then becomes

$$a = \frac{x_{1p}^2 + (x_{1p} - \overline{F})(\overline{S} - s_1 - s_2)}{x_{1p}} \qquad [443]$$

while from Eqs. [65] and [66]

$$f_1 = \frac{a(x_{1p} - \overline{F})}{x_{1p} - a} \qquad [444]$$

and

$$f_2 = \frac{a\overline{F}}{x_{1p}} . \qquad [445]$$

In order to give a practical meaning to the last three equations as well as to various unidentified equations that appear throughout this paper, readers should investigate the possibilities of Gaussian design that may exist for a system that must produce a real image of a real object when the magnification is $M' = -1$, the system focal length is $\overline{F} = 6$ units, and the distance of the image from the object is $D = 19$ units. The lens separation a (should a system of two lenses be required) must be positive. The lenses in any such system may be assumed "thin."

Details of numerous other ways in which practical application may be made of Eqs. [443] through [445] to problems that involve two separated lenses are left to the reader.

Patents Issued to RCA Inventors—Second Quarter 1984

April

G. J. Ammon Track Jump Servo System for Disc Player (4,443,869)
F. Aschwanden and W. H. Groeneweg Frequency Search System for a Phase Locked Loop (4,443,769)
R. L. Barbin Electron Beam and Deflection Yoke Alignment for Producing Convergence of Plural In-Line Beams (31,552)
E. A. Brauer Stylus Protecting Mechanism for Video Disc Player (4,442,517)
D. E. Carlson and B. F Williams Photodetector Having Enhanced Back Reflection (4,442,310)
H. Chen Electron Gun for Dynamic Beam Shape Modulation (4,443,736)
P. Datta and E. S. Polinak High Density Information Disc Lubricants (4,444,820)
A. R. Dholakia Stylus Lowering, Lifting and Cleaning Apparatus (4,443,871)
E. Dixon Digital Gyromagnetic Phase Shifter (4,445,099)
T. J. Faith, Jr. Monitor for Impurity Levels in Aluminum Deposition (4,440,799)
R. A. Gange Telephone Station Incorporating Exchange Message System (4,443,664)
L. A. Harwood and R. L. Shanley, 2nd Adjustable Coring Circuit (4,441,121)
L. A. Harwood High Voltage Protection for an Output Circuit (4,441,137)
L. V. Hedlund and J. D. Shields Tunnel for Tape Accumulation During Wind-Off (4,442,465)
A. C. Ipri and R. G. Stewart Electrically Alterable Nonvolatile Floating Gate Memory Device (4,442,447)
W. Kern Chemical Deposition Termination (4,442,134)
T. F. Kirschner Video Disc Player Having Caddy Lockout Mechanism (4,443,872)
W. L. Lehmann Television Receiver Tuning Circuit Tunable Over a Wide Frequency Range (4,442,548)
S. A. Lipp Method of Making a Focusing Color-Selection Structure for a CRT (4,443,499)
B. E. Lock Concentricity Measuring Instrument (4,439,925)
D. W. Luz Deflection Circuit With Linearity Correction (4,441,058)
R. M. Mendelson Self-Extinguishing Load Driving System (4,441,069)
C. H. Morris, Jr. Method for the Manufacture of Lapping Disc for Forming Keels on Videodisc Styli (4,440,604)
S. T. Newell Damping Mechanism for a Video Disc Stylus Holder (4,441,176)
W. R. Poff and D. W. Bartch Method for Selectively Etching Integral Cathode Substrate and Support (4,441,957)
D. R. Preslar Push-Pull Non-Complementary Transistor Amplifier (4,442,409)
G. A. Reitmeier and C. H. Strolle Compaction of Television Display Graphics in Phantom-Raster-Scanned Image Memory (4,442,545)
A. Schwarzmann Switching Microwave Integrated Bridge T Group Delay Equalizer (4,443,772)
E. B. Smith and R. A. Craft Phase Locked-Loop Generator With Closed-Loop Forcing Function Shaper (4,442,412)
K. J. Susnjara Method and Apparatus for Making Thin-Walled Plastic Articles (4,440,702)
G. A. Swartz Series Connected Solar Cells on a Single Substrate (4,443,651)
J. Tults Counter Arrangement Useful in a Frequency Locked Loop Tuning System for Measuring the Frequency of a Local Oscillator Signal (4,442,547)
H. R. Warren Crosstalk Filtering Arrangement With Variable Frequency Filtering to Remove Effects of FM Carrier (4,441,090)
D. H. Willis Degaussing Circuit for Television Receiver Having Switching Mode Power Supply (4,441,052)

May

N. Z. Assil and R. H. Hughes Color Picture Tube With Focusing Electrode Having Electrostatic Field Distortion Aperture Therein (4,449,069)

J. P. Beltz and K. W. Hang Automated Liquid Dispensing Apparatus for Spinning Surface of Uniform Thickness (4,451,507)
B. W. Beyers, Jr. Tuning Arrangement for Providing Relatively Uniform Automatic Fine Tuning Sensitivity (4,450,587)
O. H. Bismarck Speed Up Circuit (4,450,371)
F. T. D'Augustine and R. L. Berardi Apparatus and Method for Aligning the Envelope and Electron Gun Mount Assembly of a CRT (4,445,874)
F. T. D'Augustine, M. E. Terry, H. F. Welsh and F. A. Payne Automatic Mask-Frame Inserter (4,451,243)
A. G. Dingwall and V. Zazzu Switching Circuitry as for a Flash A/D Converter (4,449,118)
R. A. Dischert and K. H. Powers Transcodeable Vertically-Scanned High Definition Television System (4,449,143)
C. E. Doner Tuning Means for a Transmission Line Cavity (4,451,806)
C. W. Ebert and E. S. Thall Method of Blackening Surfaces of Steel Parts With Wet Nitrogen (4,448,612)
R. D. Faulkner, D. L. Thoman and A. F. McDonie Photomultiplier Tube Having Improved Count-Rate Stability (4,446,401)
R. D. Faulkner, D. V. Henry and D. L. Muth Broad Area Cathode Contact for a Photomultiplier Tube (4,447,758)
J. S. Fuhrer Plural Operating Mode Ambient Light Responsive Television Picture Control (4,451,849)
W. V. Fuldner Linearity Adjustment of Spacecraft Tubular Spar-Type Members (4,451,828)
P. M. Heyman Workpiece with Abraded Machine-Readable Marking Therein and Method of Making (4,446,362)
S. T. Hsu Low Resistance Contact for High Density Integrated Circuit (4,445,270)
L. M. Hughes Disc Record Player Having Stylus Cleaner (4,450,548)
K. C. Kelleher Stepper Motor Drive Circuit (4,450,394)
T. F. Kirschner Record Extraction Mechanism for Disc Player (4,451,912)
M. J. Kurina and J. Thornhill Stiffening Clamp for Self-Erecting Antenna (4,447,816)
H. G. Lewis, Jr. and S. M. Eliscu Digital Television Receiver Automatic Chroma Control System (4,447,826)
R. C. Maehl Extendible Tubular Booms for Remote Sensors (4,446,466)
A. R. Marcantonio Linear Velocity Control Means (4,450,552)
D. P. Marinelli and I. Ladany Heating Fixture (4,451,727)
J. R. Matey Apparatus and Method for Making a Video Disc (4,447,381)
K. W. McGlashan and J. R. Archer Television Raster Pincushion Distortion Correction Device (4,451,807)
J. Mount Electrical Heating Unit for Sealing Vacuum Electron Tubes (4,451,725)
E. M. Musselman Vibration Inhibiting Mesh Assembly for a Pick-Up Tube (4,446,398)
J. R. Oberman and R. G. Ott Implementation of Instruction for a Branch Which Can Cross One Page Boundary (4,449,185)
J. R. Oberman Single Chip Microcomputer With External Decoder and Memory and Internal Logic for Disabling the ROM and Relocating the RAM (4,450,524)
J. G. Ottos Method of Detecting the Vaporization of Getter Material During Manufacture of a CRT (4,445,872)
R. G. Raush Device for Supporting and Aligning Photographic Masters (4,448,522)
C. W. Reno Multi-Beam Optical Record and Playback Apparatus (4,449,212)
C. W. Reno Apparatus for Varying Track Spacing in Multi-Track Optical Record (4,449,215)
A. Schwarzmann Monolithic Reflection Phase Shifter (4,450,419)
R. L. Shanley, 2nd Picture Control for Television Receiver On-Screen Display (4,451,840)
F. Sterzer Pickup Circuit for Video Disc Including Dual-Gate FET with Injected RF (4,450,550)
C. H. Strolle Linear Interpolation Between Regularly Spaced Digital Samples (4,446,529)
J. C. Talant, 2nd Delayed Reaction Automatic Kinescope Biasing System (4,450,476)
M. Toda and S. Osaka Counterbalance System for Sagging Rotating Element (4,449,563)
I. T. Wacyk and R. G. Stewart Two Level Parity Error Correction System (4,450,562)
H. A. Weakliem and J. L. Vossen, Jr. Glow Discharge Plasma Deposition of Thin Films (4,450,787)

N. D. Welch and J. W. Woestman Circuit for Providing a Common Setpoint for Manual and Automatic Regulation of a Process Control System (4,451,879)
J. C. Whartenby, R. Brown, S. T. Rao and R. J. Menna Method for Fabricating Via Connectors Through Semiconductor Wafers (4,445,978)
D. J. Wierschke Cathode Head Having Improved Seal Means (4,445,995)
D. H. Willis Television Receiver Ferroresonant Load Power Supply (4,446,405)
J. W. Woestman System and Method for Controlling the Specific Gravity and Viscosity of the Slurry Applied to Television Picture Tube Faceplates (4,445,526)

June

L. Abbott TV Privacy System Using Gray Sync (4,454,544)
A. R. Balaban and S. A. Steckler Dual Polarity Sync Processor (4,453,183)
A. R. Balaban and S. A. Steckler Reduced Data Rate Comb Filter System (4,456,922)
V. S. Ban and E. S. Poliniak Method of Applying High-Density Information Record Lubricants (4,456,636)
D. E. Carlson and B. F Williams Photocell Utilizing a Wide-Bandgap Semiconductor Material (4,453,173)
P. Datta and E. S. Poliniak High Density Information Disc Lubricants (4,455,335)
C. B. Dieterich Coding System for Recorded Digital Audio (4,455,635)
R. A. Duschl Charge Coupled Device Based Blemish Detection System and Method (4,454,541)
R. A. Duschl Charge Coupled Device Based Inspection System and Method (4,454,545)
R. D. Faulkner and D. V. Henry Mesh Structure for a Photomultiplier Tube (4,456,852)
A. M. Goodman and R. U. Martinelli Vertical MOSFET With an Aligned Gate Electrode and Aligned Drain Shield Electrode (4,455,565)
I. Gorog Method for Evaluating Distortions in Video Disc Recordings (4,455,633)
T. F. Kirschner Video Disc Player Having Caddy Overtravel Mechanism (4,453,240)
H. P. Kleinknecht Method and Apparatus for Determining the Doping Profile in Epitaxial Layers of Semiconductors (4,456,879)
W. A. Lagoni Vertical Detail Coring Circuit to Track a Gain Adjusted Signal (4,454,533)
J. N. Laprade and R. S. Wondowski Microwave Circuit Interconnect System (4,455,537)
A. R. Moore Method and Apparatus for Determining Minority Carrier Diffusion Length in Semiconductors (4,454,472)
F. Okamoto and K. Kato Solid Particles Encapsulated With Cyanoacrylate Polymer (4,452,861)
W. T. Patton Method of Determining Excitation of Individual Elements of a Phase Array Antenna from Near-Field Data (4,453,164)
K. H. Powers Multiplier for Multiplying N-Bit Number by Quotient of an Integer Divided by an Integer Power of Two (4,455,611)
A. N. Prabhu and K. W. Hang Low Value Resistor Inks (4,452,844)
F. R. Ragland, Jr. Color Picture Tube Having Improved Temperature Compensating Support for a Mask-Frame Assembly (4,455,505)
D. J. Sauer CCD Input Circuits (31,612)
O. H. Schade, Jr. Circuitry for Generating Non-Overlapping Pulse Trains (4,456,837)
M. Toda Surface Acoustic Wave Cutterhead for Disc Recording Having a Circular Transducer (4,453,242)
J. R. Tomcavage Electron Tube Having a Low Impedance Reduced Stress Anode Structure (4,456,851)
B. Vanbreemen Projection Television Screen Having a Selected Audience Envelope (4,452,509)
R. M. Wilson RF Radial Choke for Use in Record Playback Apparatus (4,455,638)

AUTHORS

K. T. Chiang received a B.S. degree in Physics from National Tsing Hua University, Taiwan, R.O.C. in 1972, and a M.S. degree in Physics from the University of Pittsburgh in 1975. He received his Ph.D. degree in a joint program of Physics and Metallurgical Engineering from the University of Pittsburgh in 1980. His thesis research was concerned with oxidation and hot corrosion of Co-base alloys. From 1980 to 1981, he was a post-doctoral research associate in the Department of Materials and Metallurgical Engineering at the University of Pittsburgh. In 1981, he joined the Video Component and Display Division of RCA at Lancaster, PA. He is currently engaged in materials and process development of cathodes for the electron gun.

Dr. Chiang is a member of The American Society for Metals, The Metallurgical Society of AIME, and the National Association of Corrosion Engineers.

Basab B. Dasgupta received his B.Sc. and M.Sc. degrees in physics from the University of Calcutta, India and Ph.D. degree from the University of Wisconsin in 1976. He worked as a physics faculty member at two different campuses of the University of Wisconsin system and as a member of the Electrical Engineering faculty at Marquette University. His area of research has involved electronic, electromagnetic, and acoustic properties of solid surfaces. He joined RCA Consumer Electronics in 1981 as a member of the engineering staff associated with the Advanced Yoke Development group. He received a 1983 Annual Technical Excellence award from the Consumer Electronics Division.

Simon Larach received the B.S. degree in Biology and Chemistry from the City University of New York in 1943. He received the M.A. degree in Chemistry and the Ph.D. degree in Inorganic Chemistry from Princeton University in 1952 and 1955, respectively. His doctoral dissertation was titled "Magnetochemical Investigations of Luminescent Solids and Other Materials."

In 1946, he joined the Chemico-Physics group at RCA Laboratories, Princeton, NJ. He was put in charge of luminescence research in 1954, appointed Head of photoelectronic, magnetic, and dielectric research in 1956, and became Associate Laboratory Director, Materials Research Laboratory, in 1961. Dr. Larach returned to full-time research in 1967, when he was appointed a Fellow, RCA Laboratories, and a Visiting Fellow, Princeton University. Dr. Larach's research on the defect chemistry and physics of the solid state, particularly luminescent materials, and on new analytical techniques, has resulted in over 100 publications and more than 20 patents. He has contributed to the materials aspect of black-and-white as well as color TV, radar,

lasers, electroluminescence, light amplifiers, cathodes, and heaters. He directed the research culminating in the use of rare-earth, red-emitting phosphor in color television and display tubes and more recently has investigated the basic and applied aspects of phosphor adherence to kinescope-face plates.

Dr. Larach served as an ACS University-Industry Liaison Lecturer in 1968; Visiting Professor of Physics and Chemistry in 1969, and Adjunct Professor of Materials Science in 1971–74, both at Hebrew University; UNESCO Expert Consultant, Division of Technological Education and Research, Paris, 1971; Visiting Professor of Applied Physics, Ecole Polytechnique Fédérale, Lausanne, 1972; Invited NATO Lecturer on Rare Earths, Oslo, 1972; Visiting Professor of Chemistry, Princeton University, 1973; Visiting Professor of Diagnostic Radiology, Hahnemann Medical College, 1974; and, also at Hahnemann, Research Professor of Diagnostic Radiology, 1975–79. He has been Adjunct Professor of Radiology, Columbia University, College of Physicians and Surgeons, since 1979.

From 1951 to 1977, Dr. Larach won five RCA Laboratories Outstanding Achievement Awards and, in 1966, a David Sarnoff Award for Outstanding Technical Achievement, RCA's highest honor, for a team effort in science. In 1956 he was on the President's Commission on Laboratory Safety and, in 1958, he was Co-Designer, United States Solid State Exhibit, Brussel's World Fair. He is a Contributing Editor, Van Nostrand *Science Encyclopedia,* and a Divisional Editor, *Journal of the Electrochemical Society.* He is a member of the American Chemical Society, the Electrochemical Society, and Sigma Xi, and a Fellow of both the American Physical Society and the American Institute of Chemists. Dr. Larach is listed in *American Men and Women of Science* and in *Who's Who in the East.*

James J. Maley received a B.S. degree in Mathematics Education from Millersville University in 1964 and a M.S. degree in Physics from Franklin and Marshall College in 1969. Since joining RCA in 1964, he has worked on many aspects of picture tube materials and processes including internal coatings, getters, heaters, cathodes, exhaust, cathode aging, and high voltage conditioning. Since 1980, Mr. Maley has been Leader, Technical Staff, responsible for Finishing Process Development.

Mr. Maley has authored four publications and holds two U.S. patents.

Joseph E. McGowan received the Bachelor of Arts Degree with a major in chemistry from the University of Pennsylvania in 1948. He did graduate work in chemistry at Brooklyn Polytechnic Institute and Columbia University from 1950 to 1952. Mr. McGowan has been employed at the RCA Laboratories since 1965. Prior to employment at the RCA Laboratories, he worked for the Chromatic Divison of Paramount Pictures in color tube fabrication technique development. Mr. McGowan also did applications research on high voltage cables for the Okonite Co. He is currently an associate member of the technical staff and serves as a member of the Laboratories' safety committee.

Dietrich Meyerhofer studied Engineering Physics at Cornell University and received the Ph.D. degree in Physics from MIT in 1957. His thesis was a study of the ferroelectric transition in barium titanate. He has been a staff member at RCA Laboratories, Princeton, NJ, since 1958 and in 1983 he was appointed Senior Member Technical Staff. His research has been concerned with galvano-magnetic and transport properties in semiconductors and insulators, with nonlinear optical effects and the properties of infrared lasers, and with holographic processes and materials. He has carried out studies of electro-optic effects in liquid crystals and of the mechanisms underlying the functioning of dynamic scattering and field-effect display devices. For two years he was a member of an applied research group that was concerned with electronic and optic applications to the printing industry. This involved electronic photocomposition and holographic character generation, laser exposure of various kinds of printing plates, and electronic screening of black-and-white and color images for printing. The possibility of direct etching of gravure cylinders by Q-switched laser was demonstrated. The holographic investigations led to the demonstration that focussed image holograms can be used to produce high resolution microfiche. During the last few years, Dr. Meyerhofer has been concerned with testing and characterizing resists for microlithography. In the case of the positive resist systems based on diazoquinone sensitizers in novolak resins he studied the effect of exposure on the destruction of the sensitizer and the relationship between sensitizer and development rate (contrast and sensitivity). More recently, investigation of multilayer systems has become important, both for optical and for electron beam exposure. To support the research on microlithography, modeling programs were developed that allow computer simulation of the entire process. Some of the areas that have been studied are proximity printing of masks, effect of developer contrast on resist profiles, effect of resist thickness on linewidth, and the influence of sublayer properties on exposure of a bilayer system.

Dr. Meyerhofer is a member of the American Physical Society, the IEEE, and of Sigma Xi.

L. T. Sachtleben received his B.Sc. Degree in Physics and Mathematics from Antioch College. His work in optics began in the field of sound motion pictures, from which point he has continued as a generalist in optics at the RCA plants in Camden and Indianapolis. He has written or co-authored numerous published papers, and 47 patents have been inssued to him. He is a member of SMPTE, OSA, AAPT, and NCTM.

Peter David Southgate earned a B.Sc. degree in Physics and Mathematics in 1948, an M.Sc. degree in Mathematics in 1952, and a Ph.D. degree in Physics in 1958, all from London University (Imperial College). From 1948 to 1959, at Mullard Research Laboratories, England, he worked on the development of instruments, such as high-speed oscilloscopes and image-converter cameras. He also studied ultrasonic effects on metal solidification and did research on acoustic losses in semiconductors. From 1959 to 1966, at IIT

Research Institute, Chicago, IL, Dr. Southgate did research on acoustic and ultrasonic losses in insulators and semiconductors. He studied mechanisms and developed theories for charge-carrier motion, impurity diffusion, dislocation motion, and mechanical properties. During 1962–66 he headed a group working in these areas.

Dr. Southgate then joined RCA Laboratories, Princeton, NJ, as a Member, Technical Staff. His early work there included the development of an internal-ionization GaAs laser, the measurement of hot-carrier distribution functions by emitted spectra, studies of nonlinear optical coefficients (harmonic generation) in organic crystals, the development of a theory of skin-effect acoustic generation in metals, studies of mechanisms of poling in pyroelectric polymers, and the development of pyroelectric vidicon and associated equipment. Later he developed a novel method of ink-jet printing and a new type of TV-tube gun cathode. He designed and constructed prototype inspection equipment involving optical sensors and image processing, including monitors for TV-tube shadowmask apertures, for phosphor-screen line width, and for perceived mask nonuniformity. More recently he developed special equipment for detecting various types of defect in the VideoDisc, as well as algorithms to model VideoDisc stylus dynamics and stylus-disc interactions. Dr. Southgate holds nine U.S. patents, has several patents pending, and has published 34 articles in scientific journals. He is a member of the American Physical Society and has served on the Editorial Board of the *Review of Scientific Instruments*.

Harry Urkowitz received the B.S. in EE in 1948 from Drexel University and the M.S. in EE in 1954 and the Ph.D. in 1972 from the University of Pennsylvania. He was employed by the Philco Research Division from 1948 to 1964 and by the General Atronics Corp. from 1964 to 1970. Since 1970 he has been a Member of the Engineering Staff with RCA Missile and Surface Radar, Morrestown, N.J. in the Systems Engineering Department. His functions include analysis and design of radar systems with special emphasis on signal processing for detection and tracking of radar targets in noise and various forms of clutter; analysis and design of classification techniques for identification of various forms of radar target; modeling of fluctuating targets, radar clutter, and radar propagation phenomena; and the application of mathematical techniques to the stimulation of radar phenomena and of radar systems. Dr. Urkowitz is an Adjunct Professor of Electrical Engineering in Drexel University Graduate School, where he has taught courses in advanced network theory and presently teaches a course in random processes, detection, and estimation. He is the author of the book, *Signal Theory and Random Processes,* published by Artech House, Inc., in 1983.

Dr. Urkowitz is a Fellow of the IEEE and a past chairman of the Philadelphia Chapter of the Group on Circuits and Systems and the founder and a past chairman of the Philadelphia Chapter of the Group on Information Theory. He is also Chairman of the Signal Theory Subcommittee, Standards Coordinating Committee, of the IEEE Group on Information Theory. He is a member of the IEEE Aerospace and Electronic Systems Society.

Lawrence Keith White received an A.B. degree from Earlham College, Richmond, Indiana, in 1970 and a Ph.D. degree in physical chemistry from the University of Illinois, Urbana, in 1975. Dr. White was a postdoctoral research associate at the University of New Hampshire, Durham, in 1976. His doctoral and postdoctoral work dealt with the characterization of metal-ion sites in crystalline solids, metalloproteins, and biological crystals. In 1978 he joined RCA Laboratories, Princeton, NJ, as a Member of the Technical Staff. Dr. White has worked on projects dealing with IC metallization corrosion, chemical etching processes, organic dielectrics, optical modeling, and resist technology. He is currently engaged in studies of patterning technologies for advanced VLSI lithography.

Dr. White has authored over 25 technical papers and holds four US Patents. He is a member of Alpha Chi Sigma, the American Chemical Society, the Electrochemical Society, Phi Lambda Upsilon, Sigma Xi, and SPIE.