LISTING OF ENGINEERING NOTES

NUMBER	AUTHOR	SUBJECT
E-126	BEST	SCOPE SYNCHRONIZER
E-143	PERLIS	NUMERICAL SOLUTION OF LINEAR INTEGRAL EQUATIONS
E-328	ROWLAND	PULSE TRANSFORMERS AND INTERSTAGE COUPLING IN WHIRLWIND I
E-329	ADAMS	TECHNIQUES FOR USING STANDARD AUTOMATIC SUBROUTINES
E-402	HANSON	CIRCUITRY, RELAY TIMING, AND OPERATION OF WW1 TAPE OUTPUT EQUIPMENT
E-406	PAPIAN	PRELIMINARY TEST ON THE FOUR-CORE MAGNETIC-MEMORY ARRAY
E-409	FROST	RATIONAL SELECTION OF VACUUM TUBE TYPES FOR COMPUTER APPLICATIONS
E-413	EVERETT	SELECTION SYSTEMS FOR MAGNETIC CORE STORAGE
E-422	PAPIAN	RECTANGULAR-LOOP MAGNETIC CORE MATERIALS
E-449	EMERSON WERLIN	THE CADAC COMPUTER
E-451	JONES CALLAHAN	VARIATION OF TRANSISTOR COLLECTOR RESISTANCE WITH COLLECTOR VOLTAGE
E-454	BUCK	A NON-DESTRUCTIVE READ SYSTEM FOR MAGNETIC CORES
E-458-1	JEFFREY REED	THE USE OF BOOLEAN ALGEBRA IN LOGICAL DESIGN
E-458-2	JEFFREY REED	THE USE OF BOOLEAN ALGEBRA IN LOGICAL DESIGN
E-460	BUCK	THE FERROELECTRIC SWITCH
E-462	JEFFREY REED	DESIGN OF A DIGITAL COMPUTER BY BOOLEAN ALGEBRA
E-463	ECKL	INTRODUCTION TO THE THEORY OF SEMICONDUCTORS I, SOME REMARKS ON QUANTUM MECHANICS

LISTING OF ENGINEERING NOTES

NUMBER	AUTHOR	SUBJECT
E-464	BROWN	A SQUARENESS RATIO FOR COINCIDENT-CURRENT MEMORY CORES
E-468	ECKL	INTRODUCTION TO THE THEORY OF SEMICONDUCTIORS II, A BRIEF DISCUSSION OF STATISTICAL MECHANICS
E-472	MAYER	THE MIRROR: A PROPOSED SIMPLIFIED SYMBOL FOR MAGNETIC CIRCUITS
E-474	ECKL	INTRODUCTION TO THE THEORY OF SEMICONDUCTORS III, CONDUCTION IN METAL; THE FIELD-FREE CASE
E-475	BRIGGS	A MAGNETIC-CORE GATE AND ITS APPLICATION IN A STEPPING REGISTER
E-477	BUCK	MAGNETIC AND DIELECTRIC AMPLIFIERS
E-481	YOUNG	TOGGLE SWITCH INPUTS AND INDICATOR LIGHT OUTPUTS AS EXTERNAL UNITS
E-487	MAYER	NOTES ON THE LOGICAL DESIGN OF THE IBM 701 COMPUTER
E-488	GUDITZ	DELTANS IN CERAMIC ARRAY #1
E-491	MORRISON	HYSTERESIS LOOP CHARACTERISTICS OF MF-1118 FOR DIFFERENT TEMPERATURES
E-492	ECKL	INTRODUCTION TO THE THEORY OF SEMICONDUCTORS IV, QUANTUM STATES IN CRYSTALS. THE BOUND ELECTRON CASE.
E-494	NEVILLE	MODIFICATION OF THE DUMONT 304-H OSCILLOSCOPE TO REDUCE DRIFT ENCOUNTERED IN WW1 DISPLAY
E-496	HUNT	INSTRUCTIONS AND SPECIFICATIONS FOR THE MANUFACTURE OF 3:1 AND 1:1, 0.1 MICROSECOND PULSE TRANSFORMERS ON FERRITE RING CORES.
E-498	O'BRIEN	AN ELECTRONIC HOLDING CIRCUIT
E-500	KATZ GUDITZ	SWITCH-CORE ANALYSIS I
E-512	BALTZER	A METHOD FOR ACCEPTANCE TESTING OF FERRITE CORE PRODUCTION LOTS
E-523	BOYD	CORE DRIVERS -- MODEL V AND MODEL VI

LISTING OF ENGINEERING NOTES

NUMBER	AUTHOR	SUBJECT
E-525	BOYD	NORMALIZED FLIP-FLOP CHART
E-526	BOYD	HIGH-SPEED (5965) FLIP-FLOP
E-529	BUCK	MATRIX DRIVING WITH UNIDIRECTIONAL PULSES
E-531	FINE	DRIVING CURRENT MARGINS ON MEMORY TEST SETUP I*
E-533	BALTZER	EFFECT OF CURRENT PULSE DURATION ON THE PULSE RESPONSE OF M.T.C. MEMORY CORES
E-540	OLSEN PFAFF	A FAST CORE-TUBE REGISTER
E-542	HOPKINS	TIMING DIAGRAMS FOR MTC CONTROL
E-543	BOYD	FINAL SPECIFICATIONS OF THE HIGH-SPEED FLIP-FLOP
E-547	REMIS	PROPOSED HIGH INPUT-IMPEDANCE TRIGGER CIRCUITRY FOR HIGH-SPEED WWII FLIP-FLOP
E-548	CHILDRESS	PRELIMINARY REPORT - TEMPERATURE EFFECTS IN MTC-TYPE FERRITE CORES
E-550	JONES	OPEN-CIRCUIT IMPEDANCE REPRESENTATION OF TRANSISTORS*
E-561	ZIEMAN	DIFFERENTIAL VIDEO PROBE
E-563	BROWN	SPECIFICATIONS FOR A FERRITE MEMORY CORE
E-2000-1	ISRAEL	INTRODUCTION TO CODING: PART I OF II PARTS

Project Whirlwind
Servomechanisms Laboratory
Massachusetts Institute of Technology
Cambridge, Massachusetts

SUBJECT: SCOPE SYNCHRONIZER

To: 6345 Engineers, Sylvania

From: R. L. Best

Date: July 21, 1948

Description

The Scope Synchronizer will take positive pulses of any prf between 500 cycles and 6 megacycles, and divide these to any output frequency between 500 cycles and 30 kc. The output pulse may be delayed over a 10 μsec range, and synchronized with one of the input pulses if desired, to produce a jitter-free output. Two input jacks are provided, one for input prf's above 500 kc, and the other for input prf's below 500 kc.

When the input sync. selector is in position 1, the signal applied to input jack #1 is fed to the first multivibrator, which divides this frequency to about 100 kc. The free running period of this multivibrator may be varied a small amount by the "first M.V." control. The output of this first multivibrator is fed to a second multivibrator, whose free running period is controlled by the "output frequency" control.

When the input sync. selector is in position 2, the signal applied to input jack #2 is fed to the second multivibrator directly, and power supply voltage is removed from the first multivibrator.

The output of the second multivibrator triggers a single shot multivibrator, which puts out a gate, variable in length by the "delay" control from about 5-15 μsec.

When the "output sync. selector" is in position 0, smooth delay is obtained. When this switch is in position 1, the signal applied to input 1 is also fed to this delay circuit in such a manner that the trailing edge of this gate is synchronous with one of the incoming pulses, the delay then being in steps equal to the time between pulses. When this switch is in position 2, the signal applied to input 2 is used to sync. the delay.

The output circuit is an RLC peaker, which puts out a positive pulse at the end of the variable length gate. This pulse is a half sine wave of .5 μsec duration, variable in amplitude from 0 to 175 volts.

The input pulse should be at least 15 volts in amplitude. One convenient method of using the unit is to feed the "100 kc sync. pulse" from a restorer pulse generator into input 2, and put the "input sync. selector" in position 2. One mc clock pulses are fed into input 1, and the "output sync. selector" is put in position 1. Then the unit divides the 100 kc signal by some integer, which may be varied by changing the "output freq." control, and the output pulse is synchronous with a clock pulse, the delay being in steps of 1 microsecond. The "sync. gain" control varies the amplitude of the pulse fed to the multivibrator selected by the "input sync. selector", but does not vary the amplitude of the sync. pulses fed to the delay circuit.

Circuit

Drawing B-39822 shows the circuit of the Scope Synchronizer. Both multivibrators are similar except for the time constant of the crossover networks, the first multivibrator being the faster. With the "input sync. selector" on position 1, the incoming pulses are applied to the cathode of the first multivibrator through the "sync. gain" potentiometer and a crystal diode. These positive pulses are amplified by the conducting tube, and appear as positive pulses on the plate of this tube. The amplified pulses are coupled throug the crossover capacitor to the opposite grid, so that the rising exponential of the grid of the non-conducting tube has pulses superimposed on it. As this grid nears cut-off, one of these pulses causes the multivibrator to flip, whereupon the operation is repeated with the other side. Thus the multivibrator is synchronized with the applied pulses.

One plate of the first multivibrator is connected to each plate of the second multivibrator through small capacitors which differentiate the waveform of that plate of the first multivibrator. The alternately positive and negative pips which therefore appear on the plate of the off tube in the second multivibrator are passed through the crossover capacitor to the grid of the on tube, which amplifies them. These amplified pips are large enough to overide the oppositely phased pips coming directly to this plate through the small capacitor from the plate of the first multivibrator. The amplified pips are fed through the crossover capacitor to the grid of the off tube, being superimposed upon this rising exponential. When this grid nears cutoff, it is one of these pips which causes the multivibrator to flip, synchronizing it. The operation is then repeated for the opposite side of the multivibrator. When the "input sync. selector" is in position 2, the second multivibrator operates in a manner similar to the preceding description of the first multivibrator, the sync. pulses being applied to the cathode.

A small capacitor is connected from one plate of the second multivibrator to a voltage divider, differentiating this waveform, and putting the result at a d-c potential of about 20 volts. Only the negative pip caused by this plate conducting is large enough to pass through the crystal connected to the normally on grid of the single shot or delay multivibrator, and this pip flips this delay multivibrator. The "delay" control determines the length of the gate generated at the plate of this tube. The normally off plate conducts during the delay interval, and is connected to the output sync. selector through a crystal diode and capacitor. The pulses selected by the "output sync. selector" are fed to this plate, and to the exponentially rising grid of the off triode through the crossover capacitor. The end of the gate is therefore synchronous with the pulses selected by the "output sync. selector" switch. After the delay multivibrator has returned to its original state, the crystal diode connected to the normally off plate passes no synchronizing pulses, since its cathode is at +150 volts, while its anode is at about +70 volts.

The output circuit is an RLC peaker, operating on the positive gate generated by the normally on plate of the delay multivibrator. This gate is R-C coupled to a variable bias, varied by the "output amplitude" control, which determines how heavily the output tube conducts during the gate interval. At the end of the gate interval, the output triodes are sharply cut off, and the resulting inductive voltage swing in the plate circuit is the output pulse. Negative voltage swings in the plate circuit are clipped by crystal diodes. Due to the method of coupling the gate to the peaker, variations in output frequency and delay effect the output amplitude. The output pulse is a $\frac{1}{2}$ μsec half sine wave, that may be varied from zero to + 175 volts.

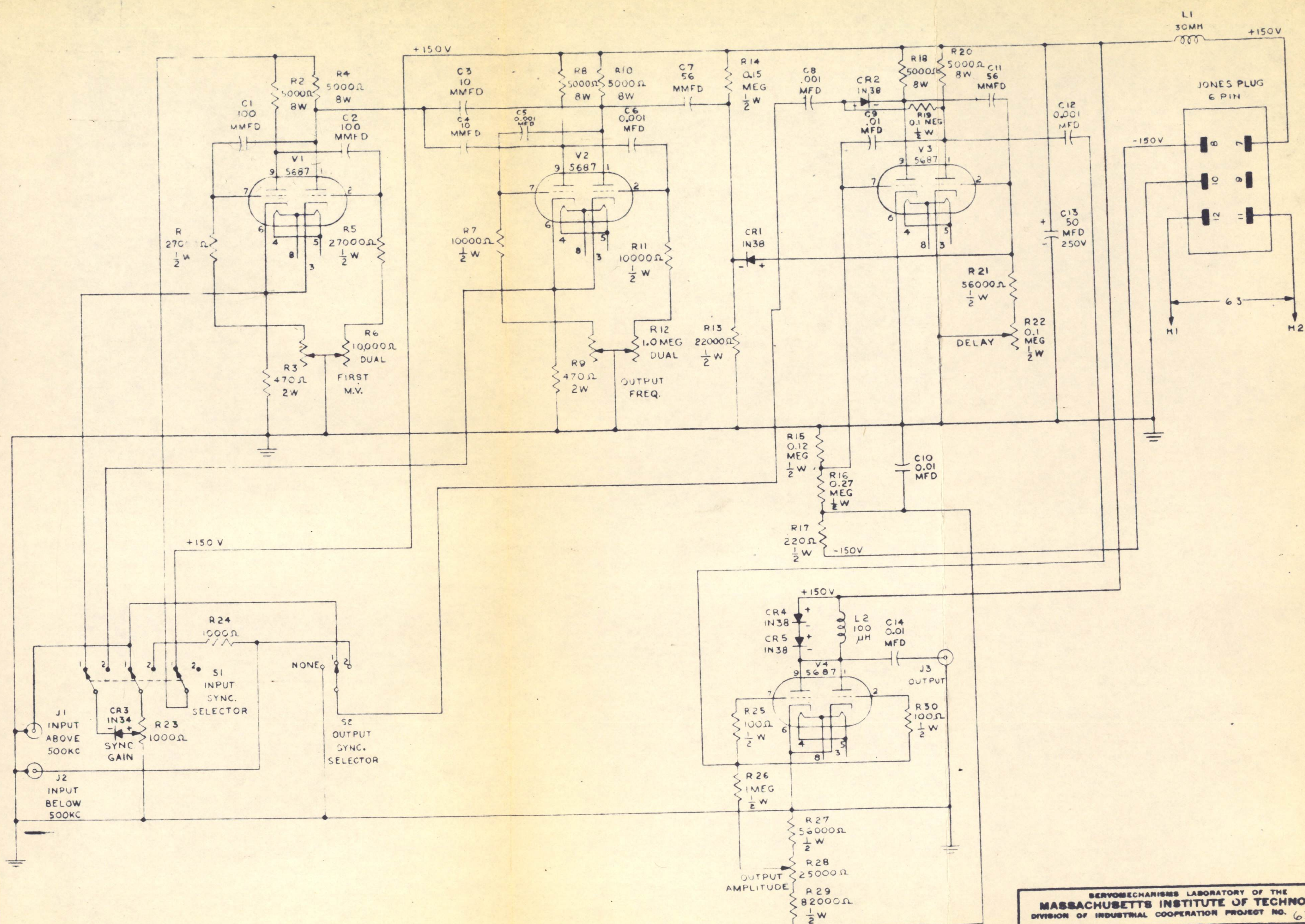
Signed R. L. Best
R. L. Best

Approved N H Taylor
N. H. Taylor

RLB/spr

Drawing: 39822

B-39822



SERVOMECHANISMS LABORATORY OF THE
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
DIVISION OF INDUSTRIAL COOPERATION PROJECT NO. 6345

SCOPE SYNCHRONIZER

SCALE:	DR A.M.G.	B-39822
ENG. R.L.B.	CK Y.L. 6-21-48	APP. B-REDUCTION

USE IN 6345 MEMO E126

Project Whirlwind
Servomechanisms Laboratory
Massachusetts Institute of Technology
Cambridge, Massachusetts

SUBJECT: NUMERICAL SOLUTION OF LINEAR INTEGRAL EQUATIONS

To: Mathematics Group

From: Alan J. Perlis

Date: September 16, 1948

Introduction:

The purpose of this report is a survey of numerical methods for solving integral equations. Only linear integral equations of real variables will be considered. There are three important types of such equations:

(a) The Fredholm equations of the second kind

$$f(x) = g(x) + \lambda \int_a^b K(x,y) f(y) dy \quad \text{(non-homogeneous)}$$

$$f(x) = \lambda \int_a^b K(x,y) f(y) dy \quad \text{(homogeneous)}$$

(b) The Volterra equation

$$f(x) = g(x) + \int_a^x K(x,y) f(y) dy$$

and

(c) The equation of the first kind

$$g(x) = \int_a^b K(x,y) f(y) dy$$

In each of the above equations the function f(x) is to be obtained from the given functions g(x) and K(x,y). Under quite general conditions (a), and (b) possess unique solutions whereas (c) will, in general, have many solutions except under certain restrictive constraints.

Thus far, at least by comparison, say, with linear algebraic equations and differential equations, little has been accomplished in developing numerical methods for solving integral equations. The superiority over differential equations in expressing many physical problems has only recently been exploited.

In general, it would be nicest to attack each equation by a method best suited to its own structure. This puts a premium on analysis. The attitude taken here will be to develop quite general numerical methods which can most easily be treated by a computer of the digital type.

On the same level of importance as the design of a solving method for an equation is an appreciation of the status of errors in each step of the problem. Besides the usual numerical round-off and truncation errors, all methods of solution introduce an important stability error caused by the lapse of a transcendental problem into a discrete one. In general, the numerical methods show excellent stability (at least where unique solutions to the problem exist) due chiefly to the smoothing properties of integration.

2. The Linear Integral Equation of the First Kind

The equation to be considered is

$$f(x) = \int_a^b K(x,y)\, g(y)\, dy \qquad (1)$$

with the functions f(x) and K(x,y) given and g(y) to be determined. Examples of this type of equation are:

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{iyx} g(y)\, dy \qquad \text{(The Fourier transform)}$$

and

$$f(x) = \int_0^{\infty} e^{-xy} g(y)\, dy \qquad \text{(The Laplace transform)}$$

Often the solution g(y) will not be unique; for, to the solution of (1) must be added solutions, if they exist, of

$$0 = \int_a^b K(x,y)\, g_i(y)\, dy \qquad (i = 1, 2, \ldots)$$

In both of the above mentioned examples definite inversion formulae of the type:

$$g(y) = \int_L h(x,y) f(x)\, dx \qquad \text{(L, a path in the x plane)}$$

yield analytic expressions for g(y). In general, such inversion formulae do not exist. The integral equation does not always have solutions, since for a given K(x,y) f(x) must possess certain properties dependent upon the kernel. If K(x,y) is a polynomial in x, then f(x) must be a polynomial in x. Likewise if K(a,y) = 0 for all y and if the equation is satisfied for all y, then f(a) = 0. The function f(x) is not perfectly arbitrary, for if K(x,y) is a polynomial in x, then f(x) must be a polynomial. That is, we must restrict f(x) to a class of functions associated with K(x,y). A theorem to this effect will be discussed later. In certain special cases an equation of the first kind can be reduced to one of the second kind and here f(x) can be highly arbitrary.

As regards the uniqueness of the solution g(y), where g(y) is constrained to be continuous, the following may be said. If the kernel, K(x,y), is "definite", i.e.

$$\int_a^b\!\!\int_a^b K(x,y)\, w(x)\, w(y)\, dx\, dy > 0$$

for all continuous w(x), then the equation

$$\int_a^b K(x,y)\, w(y)\, dy = 0$$

has no continuous solution w(y) except w(y) = 0. In particular, if two continuous solutions $g_1(y)$ and $g_2(y)$ of the equation exist, then, for K(x,y) definite, $g_1(y) = g_2(y)$; i.e. the solution is unique. In general a given K(x,y) will not be definite but may be made so by the following artifice. Form:

$$L(x,y) = \int_a^b K(x,t)\,K(y,t)\,dt.$$

Since L(x,y) is symmetric its characteric values are real. Furthermore, if $\phi_n(x)$ is a characteristic function,

$$\phi_n(x) = \lambda_n \int_a^b L(x,y)\,\phi_n(y)\,dy$$

From which it follows that $\lambda_n \geq 0$ since

$$\int_a^b \phi_n^2(x)\,dx = \lambda_n \int_a^b \phi_n(x)\,dx \int_a^b L(x,y)\,\phi_n(y)\,dy$$

and

$$= \lambda_n \int_a^b \phi_n(x)\,dx \int_a^b \phi_n(y)\,dy \int_a^b K(x,t)K(y,t)\,dt$$

$$= \lambda_n \int_a^b \left(\int_a^b K(x,t)\,\phi_n(t)\,dt \right)^2 dx$$

hence $\lambda_n \geq 0$. Indeed if the $\{\phi_n(x)\}$ form a complete set,

$$\int_a^b K(x,t)\,\phi_n(x)\,dx \neq 0 \quad (n = 1, 2, \ldots)$$

so that $\lambda_n > 0$. Hence L(x,y) is definite and the integral equation:

$$u(x) = \int_a^b L(x,y)\,g(y)\,dy$$

where

$$u(x) = \int_a^b f(t)\,K(x,t)\,dt$$

has a unique continuous solution. If the function g(y) is restricted to be continuous f(x) will usually be required to satisfy linear conditions of the form:

$$\int_a^b w(x)\,f(x)\,dx = 0, \quad \int_a^b w(x)\,K(x,y)\,dx = 0$$

An analogy to linear algebraic equations proves fruitful. The equation is analogous to the system:

$$f_x = \sum_{y=1}^{n} K_{x,y}\,g_y \quad (x = 1, 2, 3, \ldots, m)$$

and can be obtained from the system in the limit as $m, n \to \infty$. The properties of the system depend on the relation between m and n. Hence three cases can occur:

(1) $m > n$ implies the existence of relations of the form

$$\sum_{x=1}^{x=m} a_x K_{x,y} = 0 \qquad \text{for all } y$$

and then consistency will demand that f_x satisfy:

$$\sum_{x=1}^{x=m} a_x f_x = 0$$

In the limit these relations may take several forms, an obvious one being:

$$\int_a^b a(x)\, f(x)\, dx = 0$$

(2) $m = n$. A unique set of g_ys exists and is given by:

$$g_y = \sum_{x=1}^{x=m} a_x K_{x,y} \qquad y = (1, 2, \ldots, m)$$

provided no relation

$$\sum_{x=1}^{x=m} a_x K_{x,y} = 0 \qquad \text{(all } y)$$

is satisfied.

(3) $m < n$ yields an infinite number of sets of solutions but particular sets may be obtained by imposing $(n - m)$ linear conditions on g_y.

Hence, in general, it is too much to expect that the solution will be unique. An important question, particularly in numerical calculations, is the range in x over which f(x) is given. Decreasing the range may well involve making g(y) zero over a portion of the range to maintain uniqueness. In the particular case of the moment problem of Stieltjes, f(x) is defined only over an enumerable set of values $\{x_i\}_{i=1}^{i=\infty}$ and the side conditions induce a unique g(y).

In general, then, if the conditions are satisfied which will make the solution unique we shall expect to obtain an expression of the form:

$$g(y) = H_K f(y)$$

where H_K may be built up of linear combinations of the type

$$p(y)\frac{d^n f(y)}{dy}, \quad f(y+\beta), \quad \int_a^b H(y,x)\, f(x)\, dx$$

In two cases mentioned the inversion formula is of the highly desirable form:

$$g(y) = \int_L H(y,x)\, f(x)\, dx \qquad \text{(L some path)}$$

Only rarely will this integral be of the same type as that defining f(x). For a g(y) may be discontinuous at some $y = y_o$, and still yield a continuous f(x), hence the integral for g(y) must be improper, that is H(y,x) must be discontinuous or the limits be infinite.

In those cases where the equations of condition are not satisfied it will then not be possible to obtain an exact representation of f(x) by an integral. But by analogy to the method of least squares, the problem of solving the integral equation is essentially equivalent to that of minimizing an integral, i.e. a problem in the calculus of variations. The point being to obtain the best possible approximation to f(x) in the least squares sense by an integral of the given type. As in the least squares procedure we may introduce a positive weight function q(x) which weights f(x) according to some law. Then if

$$f(x) = \int_a^b K(x,y)\, g(y)\, dy$$

$$\int_a^b q(x) K(x,s) f(x)\, dx = \int_a^b q(x) K(x,s)\, dx \int_a^b K(x,y) g(y)\, dy$$

or

$$\phi(s) = \int_a^b L(s,y)\, g(y)\, dy$$

where

$$L(s,y) = \int_a^b K(x,s) K(x,y)\, q(x)\, dx$$

and

$$\phi(s) = \int_a^b q(x) K(x,s) f(x)\, dx.$$

If we assume that k(x,y) is "perfect", i.e.

$$\int_a^b K(x,y)\, u(y)\, dy = 0$$

implies, if u(y) is continuous, that u(y) = 0. The L(s,y) is definite, and symmetric. Hence the solution is unique, but may not possess a continuous solution g(y). This can be answered by ascertaining the existence or non-existence of a continuous function g(y) which minimizes

$$W = \int_a^b q(x) \left(f(x) - \int_a^b K(x,y)\, g(y)\, dy \right)^2 dx$$

over a set of values in x. This minimum may not always exist but W may be made as small as desired, and thus an approximate representation for f(x) is obtained.

A rewarding way of examining the integral equation of the first kind is as an operator equation:

$$f(x) = \int_a^b K(x,y) g(y) dy \quad \text{as} \quad f(s) = L_K\, g(s)$$

and then the problem is to find L^{-1}_k such that

$$g(s) = L_K^{-1} f(s)$$

and, in general, will not be unique. For numerical work the problem of uniqueness becomes largely an academic one, in any case. The interesting question here is under what general conditions can a solution be found; how accurate will it be; and is there a quite general, easily arithmetized algorithm which will yield a reasonably accurate result. A general theorem by Bateman proves useful.

The equation is:

$$f(x) = \int_a^b K(x,y)\, g(y)\, dy.$$

Either K(x,y) is symmetric or it is symmetrized by the operation:

$$u(s) = \int_a^b f(x) K(s,x)\, dx = \int_a^b K(s,x)\, dx \int_a^b K(x,y) g(y)\, dy$$

$$u(s) = \int_a^b K(s,x) K(x,y)\, dx \int_c^b g(y)\, dy = \int_a^b L(s,y) g(y)\, dy$$

Henceforth K(x,y) will be assumed to be symmetric. The theorem then is:

(i) If K(x,y) has a complete set of characteristic functions, $\phi_n(x)$, satisfying

$$\phi_n(x) = \lambda_n \int_a^b K(x,y)\, \phi_n(y)\, dy$$

and

(ii) f(x) can be expanded in an absolutely and uniformly convergent "Fourier" series

$$f(x) = \sum_{n=0}^{\infty} C_n \phi_n(x)$$

then there exists a g(y) such that

$$\left| f(x) - \int_a^b K(x,y)\, g(y)\, dy \right|$$

may be less than any arbitrary $\epsilon > 0$. The proof of the theorem provides the algorithm for the solution.

Proof:

Define: $$F(z,x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} z^{2n+1} L_K^{2n+1} f(x) \quad ; \quad L_K f(x) = \int_a^b K(x,y) f(y)\, dy$$

and

$$H(z,x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} z^{2n} L_K^{2n} f(x)$$

The series are abs. convergent for finite x. For:

Let $|f(x)| < \tilde{f}$

$|K(x,y)| < \tilde{K}$

They imply

$$|L_K^n f(x)| < (b-a)^n \tilde{f} \tilde{K}^n$$

and so each series becomes less than the exponential series.

Pick an M and write:

$$g(y) = 2 \int_0^M F(y,z)\,dz$$

$$\int_a^b K(x,y) g(y)\,dy = 2 \int_a^b K(x,y)\,dy \int_0^M F(y,z)\,dz$$

$$= 2 \int_0^M dz \int_a^b F(y,z) K(x,y)\,dy$$

Integration term by term yields

$$2 \int_a^b K(x,y) F(y,z)\,dy = 2 \sum_{n=0}^{\infty} (-1)^n \frac{z^{2n+1}}{n!} L_K^{2n+2} f(x)$$

$$= -\frac{d}{dz} H(z,x)$$

Therefore:

$$\int_a^b K(x,y) g(y)\,dy = -\int_0^M \frac{d}{dz} H(z,x)\,dz$$

$$= f(x) - H(x,M)$$

An expression for large M, of H(x,M), is to be found.

Use: $f(x) = \sum_{n=1}^{\infty} c_n \phi_n(x)$

Form: $\int_a^b K(x,y) f(y)\,dy = \sum_{n=1}^{\infty} \frac{c_n}{\lambda_n} \phi_n(x)$ which converges.

Also:

$$L_K^r f(x) = \sum_{n=1}^{\infty} \frac{c_n}{\lambda_n^r} \phi_n(x) \qquad (r = 1, 2, \ldots)$$

So:

$$H(x,z) = \sum_{n=0}^{\infty}(-1)^n \frac{Z^{2n}}{n!} L_K^{2n} f(x) = \sum_{n=0}^{\infty}(-1)^n \frac{Z^{2n}}{n!}\left(\sum_{n=1}^{\infty} \frac{c_n}{\lambda_n^{2n}} \phi_n(x)\right)$$

Now: $\sum_{n=m}^{\infty} |c_n \phi_n(x)| < \epsilon$ implies, for m large enough,

$$\sum_{n=m}^{\infty} \left| \frac{c_n}{\lambda_m^{2n}} \phi_n(x)\right| < \frac{\epsilon}{|\lambda_m^{2n}|}$$

since $\lambda_{n+1} > \lambda_{n+2}$ (by formal ordering of the eigenfunctions)

Hence:

$$H(x,z) = \sum_{n=0}^{\infty}(-1)^n \frac{Z^{2n}}{n!}\left(\sum_{n=1}^{m} \frac{c_n \phi_n(x)}{\lambda_n^{2n}}\right) + \delta$$

where

$$|\delta| < \epsilon e^{-Z^2/\lambda_m^2}$$

So rearranging:

$$H(x,z) = \sum_{n=1}^{m} c_n \phi_n(x) \cdot e^{-Z^2/\lambda_n^2} + \delta.$$

Finally:

$$H(x,z) = \sum_{n=1}^{\infty} c_n \phi_n(x) e^{-Z^2/\lambda_n^2}$$

so now

$$\left|\sum_{n=p}^{\infty} c_n e^{-z^2/\lambda_n^2} \phi_n(x)\right| < \left|\sum_{n=p}^{\infty} c_n \phi_n(x)\right| < \epsilon/2$$

and $\left|\sum_{n=1}^{p} c_n e^{-M^2/\lambda_n^2} \phi_n(x)\right| < \epsilon/2$, for M large enough.

Therefore $|H(x,M)| < \epsilon.$

Hence $\left|\int_a^b K(x,y)\, g(y)\, dy - f(x)\right| < \epsilon$

and so the solution is

$$g(y) \sim 2\int_0^M F(y,z)\, dz$$

If, in addition, g(y) is constrained to remain finite over the entire range of integration, more can be said. It has been shown that

$$\left|\int_a^b K(x,y)\, g(y)\, dy - f(x)\right| < \epsilon$$

but not that

$$\lim_{\epsilon \to 0} g_\epsilon(y) = g(y)$$

If, in addition to the finiteness condition on g(y) the "derived series" $\sum |\lambda_n \varsigma_n \tilde{\phi}_n|$ is convergent, $\tilde{\phi}_n = |\phi_n|_{max}$ on (a,b) then there exists a g(y) defined by $g(y) = 2\int_0^\infty F(y,z)\,dz$

such that

$$f(x) = \int_a^b K(x,y)\, g(y)\, dy$$

where F(y,z) is defined by the series on page (7).

As an example of this method of solution, consider the equation:

$$e^{-s^2} = \int_0^\infty g(t) \cos st \, dt$$

The solution is known to be $g(t) = \frac{1}{\sqrt{\pi}} e^{-t^2/4}$

The function F(t,x) is formed:

$$F(t,x) = x\int_0^\infty e^{-s^2} \cos st \, ds - x^3 \int_0^\infty \cos ts_1\, ds_1 \int_0^\infty \cos s_1 s_2\, ds_2 \int_0^\infty e^{-s_3^2} \cos s_3 s_2\, ds_3 + \cdots\cdots$$

$$= e^{-t^2/4} \sum_{n=0}^\infty \frac{(-1)^n}{n!} \frac{(\sqrt{\pi})^{2n+1}}{2^{n+1}} x^{2n+1}$$

$$g(t) = 2\int_0^\infty e^{-t^2/4} \left(\sum_{n=0}^\infty \frac{(-1)^n}{n!} \frac{(\sqrt{\pi})^{2n+1}}{2^{n+1}} x^{2n+1} \right) dx$$

$$= \frac{1}{\sqrt{\pi}} e^{-t^2/4} \int_0^\infty d\left(e^{-\pi/2\, x^2}\right)$$

$$= \frac{1}{\sqrt{\pi}} e^{-t^2/4}$$

3. Solution in Terms of Orthogonal Functions:

The conditions for the existence of a solution of the equation of the first kind may be rather neatly stated in terms of orthogonal functions. The treatment will be restricted to equations and functions satisfying the following conditions:

(i) The variable y is real on (a,b); the variable x is real on the same finite interval of definition (a,b).

(ii) The functions f(x), k(x,y) are real.

(iii) The kernel k(x,y) is bounded and integrable squared (L^2) in x and y; the function f(x) is integrable squared (L^2) in (a,b). Hence g(y) is L^2 in (a,b).

It will be convenient to say that if

$$\int_a^b (f(y)-g(y))^2\,dy = 0$$

Then f(y) = g(y).

A set of real functions $\{\phi_n(x)\}$ are ortho-normal L^2 on (a,b) if they satisfy:

$$\int_a^b \phi_n(x)\,\phi_m(x)\,dx = \delta_{mn} \quad , m,n=1,2,3,\ldots$$

If the set is orthogonal and L^2 on (a,b) it may be made orthonormal. The set is said to be complete if no function u(x) in L^2 exists for which

$$\int_a^b \phi_n(x)\,u(x)\,dx = 0 \quad (\text{all } n)$$

If the set $\{\phi_n(x)\}$ is not complete, it may be completed by adjoining to it a set $\{\phi_n'(x)\}$ (13) which is called the complementary set to $\{\phi_n(x)\}$.

Bessel's inequality is:

$$\sum_{n=1}^{\infty}\left(\int_a^b f(x)\,\phi_n(x)\,dx\right)^2 = \sum_{n=1}^{\infty} f_n^2 \le \int_a^b f(x)^2\,dx$$

If the set $\{\phi_n(x)\}$ is complete the equality holds. In addition, for a complete set, the relation

$$\int_a^b f(x)\,g(x)\,dx = \sum_{n=1}^{\infty} f_n\,g_n$$

is valid. If not commlete the equality holds if and only if

$$\int_a^b g(x)\,\phi_n(x)\,dx = 0 \qquad n = 1, 2$$

implies

$$\int_a^b f(x)\,g(x)\,dx = 0$$

for all $g(x)$ on $L^2(a,b)$. This can be put another way. In order that $f(x)$ be orthogonal to all $g(x)$ which satisfy the above, it is sufficient that $f(x)$ be orthogonal to the set $\{\phi_n'(x)\}$ complementary to $\{\phi_n(x)\}$.

The fundamental Riesz-Fischer theorem states: if $\{f_n\}_{n=1}^{\infty}$ exists such that $\sum_1^{\infty} f_n^2 < \infty$ then there exists and $f(x)$ in $L^2(a,b)$ such that

$$f_n = \int_a^b f(x)\,\phi_n(x)\,dx \qquad n = 1, 2, \ldots$$

where

$$\int_a^b \phi_m(x)\,\phi_n(x)\,dx = 0 \qquad m \neq n = 1, 2, \ldots$$

$$\int_a^b \phi_n^2(x)\,dx < \infty \qquad n = 1, 2, \ldots$$

and if the set $\{\phi_n(x)\}$ is complete, $f(x)$ is unique except possibly for a set of measure zero.

One point which may now be made clear is that there functions for which the equation (1) has no solution. For if $f(x)$ L^2 on some (a,b), then for some orthogonal set on (a,b);

$$c_n = \int_a^b f(x)\,\Theta_n(x)\,dx = \int_a^b A_n(y)\,g(y)\,dy$$

where

$$A_n(y) = \int_a^b K(x,y)\,\Theta_n(x)\,dx$$

Then

$$c_n^2 \leq \int_a^b A_n^2(y)\,dy \int_a^b g(y)^2\,dy.$$

Let

$$\int_a^b A_n^2(y)\,dy = a_n^2$$

so that $\frac{c_n^2}{a_n^2}$ be bounded is a necessary condition if $f(x)$ is to be a solution of (1); however it is not sufficient. But:

$$\sum_{n=1}^{N} a_n^2 = \int_a^b \Big[\sum_{n=1}^{N} A_n^2(y)\Big]dy \leq \int_a^b\int_a^b K^2(x,y)\,dy\,dx.$$

so that $\sum_{n=1}^{\infty} a_n^2$ is convergent and furthermore depends only on the kernel. But there are series $\sum_{n=1}^{\infty} c_n^2$ and $\sum_{n=1}^{\infty} a_n^2$ which converge, and yet for which $\frac{c_n^2}{a_n^2}$ is unbounded. For example: $\frac{c_n^2}{a_n^2} = n$ and

$$\sum_{n=1}^{\infty} c_n^2 = \sum_{n=1}^{\infty} n a_n^2 = \sum_{n=1}^{\infty} \frac{a_n'^2}{n^3}$$ i.e. both series converge.

It is possible to reduce the solution of the given equation to that of the following problem: Being given a denumerable set of constants $\{c_n\}$ and a set of functions $\{A_n(y)\}$, to find a g(y) which satisfies:

$$c_n = \int_a^b A_n(y)\, g(y)\, dy \qquad n = 1, 2, \ldots$$

The following may be said at once: If the $A_n(y)$ form an ortho-normal set on (a,b) the necessary and sufficient condition for a solution g(y) is that

$$\sum_{n=1}^{\infty} c_n^2 < \infty$$

then the solution is the limit, in the mean, of

$$\sum_{n=1}^{N} c_n A_n(y).$$

Naturally, if $\{A_n(y)\}$ be not closed there must be added, to the above, solutions of:

$$\int_a^b A_n(y)\, g(y)\, dy = 0$$

If the $\{A_n(y)\}$ are not ortho-normal, they may be made linearly independent by suppressing those which obey relations of the form:

$$A_n(y) = \sum_{i=1}^{n-1} \gamma_i^{(n)} A_i(y)$$

This creates similar equations among the c's:

$$c_n = \sum_{i=1}^{n-1} \gamma_i^{(n)} c_i.$$

The remaining equations are ortho-normalized. Hence one gets series of the form

$$\sum_{n=1}^{\infty} b_n^2 = \sum_{n=1}^{\infty} (a_1^{(n)} c_1 + \cdots \; a_n^{(1)} c_n)$$

which must converge and the solution is the limit in the mean of

$$\sum_{n=1}^{N} b_n B_n(y).$$

where $\{B_n(y)\}$ are the derived ortho-normal set. Conversely if this problem is solved, where the C_n are the Fourier coefficients of f(x), the equation is solved.

The solutions previously given demand knowledge of the characteristic values and functions of the kernel or of the symmetrized definite kernel constructed from it.

It will now be show that solving the integral equation is equivalent to that of maximizing a functional equation whose maximum value, if attained, is the solution. The conditions that the maximum be attained, then, are the conditions previously obtained which insure the solutions.

The functional H(g) is defined as

$$H[g] = \frac{\left(\int_a^b f(x)\,\tilde{g}(x)\,dx \right)^2}{\int_a^b \int_a^b L(x,y)\,\tilde{g}(x)\,\tilde{g}(y)\,dx\,dy}$$

where L(x,y) is the previously defined symmetrized kernel:

$$L(x,y) = \int_a^b K(x,t)\,K(y,t)\,dt;$$

f(x) is defined by:

$$f(x) = \int_a^b K(x,y)\,g(y)\,dy.$$

and g(x) is any $L^2(a,b)$ function. The denominator of $H[g]$ is bounded away from zero if L(x,y) is "closed" or if f(x) is orthogonal to each member of the complimentary set of eigen functions of L(x,y). This will be assumed. Define now the "Fourier" coefficients:

$$c_i = \int_a^b f(x)\,\phi_i(x)\,dx$$

$$\tilde{g}_i = \int_a^b \tilde{g}(x)\,\phi_i(x)\,dx$$

where the $\{\phi_i(x)\}$ are the characteristic functions of L(x,y). Then $H[g]$ can be written as:

$$H[g] = \frac{\left(\sum_{n=1}^{\infty} \tilde{g}_n f_n \right)^2}{\sum_{n=1}^{\infty} \frac{\tilde{g}_n^2}{\lambda_n^2}}$$

Now introduce the identity:

$$\sum_{i=1}^{N} a_i^2 \sum_{i=1}^{N} b_i^2 = \left(\sum_{n=1}^{N} a_i b_i\right)^2 + \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{i}(a_i b_j - a_j b_i)^2$$

Now as $N \rightarrow \infty$, if $\sum_{n=1}^{\infty} a_i^2$ and $\sum_{n=1}^{\infty} b_i^2$ converge then $\sum_{n=1}^{\infty} a_i b_i$ converges and hence the series

$$\sum_{i=1}^{\infty}\sum_{j=1}^{i}(a_i b_j - a_j b_i)^2$$

converges. The substitution $a_i = \lambda_i c_i$, $b_i = \tilde{g}_i/\lambda_i$ yields, the inequality

$$\sum_{i=1}^{N} \lambda_i^2 c_i^2 \geq \frac{\left(\sum_{i=1}^{N} \tilde{g}_i c_i\right)^2}{\sum_{i=1}^{N} \tilde{g}_i^2/\lambda_i^2}$$

Now let $N \rightarrow \infty$; the right side approaches $H(g)$ and if the series converges $H(g)$ is bounded. Furthermore $\sum_{i=1}^{\infty} \tilde{g}_i^2/\lambda_i^2$ converges so that the identity can be written:

$$H(g) = \sum_{i=1}^{\infty} \lambda_i^2 c_i^2 - \frac{\frac{1}{2}}{\left(\sum_{i=1}^{\infty} \tilde{g}_i^2/\lambda_i^2\right)} \sum_{i=1}^{\infty}\sum_{j=1}^{i}\left(\tilde{g}_i c_j \frac{\lambda_j}{\lambda_i} - \tilde{g}_j c_i \frac{\lambda_i}{\lambda_j}\right)^2$$

Now put $g_i = \lambda_i^2 c_i$. Hence for the function $g(x)$ possessing these Fourier coefficients $H(g)$ attains the maximum. Hence if the series $\sum_{i=1}^{\infty} \lambda_i^2 c_i^2$ converges and $\sum_{i=1}^{\infty} \lambda_i^4 c_i^2$ converges, then there exists a $\tilde{g}(x)$ whose Fourier coefficients are $\lambda_i^2 c_i$ and which maximizes the functional $H(g)$; furthermore, this $\tilde{g}(x)$ satisfies the integral equation (1).

For numerical work, one would define $\tilde{g}_i = \lambda_i^2 c_i$ for $i \leq N$ and $\tilde{g}_i = 0$ for $i > N$. The N being selected as large as necessary for a good approximation. The method used for the variational procedure would be, perhaps the Trefftz variation of the Rayleigh Ritz method. But here again a knowledge of the characteristic values of $L(x,y)$ are necessary.

4. The Method of Steepest Descents

The equation to be solved is:

$$f(x) + \int_a^b K(x,y)\,g(y)\,dy = 0$$

We form L (x,y):

$$\int_a^b f(x)\,K(x,y)\,dx = \int_a^b K(x,y)\,dx \int_a^b K(x,t)\,g(t)\,dt$$

$$L(y,t) = \int_a^b K(x,y)\,K(x,t)\,dx$$

$$u(y) = \int_a^b L(y,t)\,g(t)\,dt = -v(y).$$

So we wish to find a g (t) such that

$$\int_a^b L(y,t)\,g(t)\,dt + v(y) = 0$$

and the solution of this problem is the same as of (1).

The method of steepest descents can be used as an iterative method providing certain conditions are satisfied. It will be assumed that L(x,y) is such that, for all v(y) under consideration,:

(I) $$\int_a^b v(y)\,dy \int_a^b L(y,t)\,v(t)\,dt \geq m \int_a^b v^2(t)\,dt \quad ; m > 0$$

(II) $$\int_a^b \left\{ \int_a^b L(y,t)\,v(t)\,dt \right\}^2 dy \leq M \int_a^b v^2(y)\,dy \quad ; M > 0.$$

(I) can be written as:

$$\int_a^b dx \left(\int_a^b K(x,y)\,v(y)\,dy \right)^2 \geq m \int_a^b v^2(t)\,dt$$

Now if the characteristic functions of L (y,t) form a closed set, an m can be found such that (I) holds true for all v(y) which are $L^2(a,b)$. If the characteristic functions of L (y,t) do not form a closed set, a set $\{\phi'(y)\}$ is adjoined to yield the complete set $\{\Phi(y)\} = \{\phi(y)\} + \{\phi'(y)\}$ Then for there to exist an m, must all v(y) be orthogonal to the complementary set $\{\phi'(y)\}$.

Apply the Schwartz inequality, now, to (II) and obtain:

$$\int_a^b dt \left| \int_a^b K(y,t) v(y) dy \right|^2 \leq \int_a^b dt \int_a^b |L(y,t)|^2 dy \int_a^b v(y)^2 dy$$

$$\leq \int_a^b v^2(y) dy \int_a^b \int_a^b (L(y,t))^2 dy\, dt$$

$$\leq M \int_a^b v^2(y) dy.$$

So (II) follows if the kernel L (y,t) is $L^2(a,b)$. As in the case of algebraic equations the functional W(g), defined by

$$W(g) = \tfrac{1}{2} \int_a^b g(y) dy \int_a^b L(y,t) g(t) dt + \int_a^b v(y) g(y) dy$$

is considered. Squaring

$$\int_a^b g(y) dy \int_a^b L(y,t) g(t) dt + \int_a^b v(y) g(y) dy.$$

yields the inequality:

$$\left\{ 2 \int_a^b g(y) dy \int_a^b L(y,t) g(t) dt \right\} W(g) + \left\{ \int_a^b v(y) g(y) dy \right\}^2 \geq 0.$$

Schwartz's inequality applied to the second integral and condition (I) applied to the first give:

$$2m \left(\int_a^b g^2(y) dy \right) W(g) + \int_a^b g^2(y) dy \int_a^b v^2(y) dy \geq 0$$

Hence: $W(g) \geq -\frac{1}{m} \int_a v^2(y) dy$

and W(g) has finite lower bound independant of g(y).

Now, to improve and approximation, $g^{(v)}(y)$, we consider the functional:

$$W(g^{(v+1)}) = W(g^{(v)} + Y w(y))$$

where Y, a real number, and w(y), a function, are to be determined.

Then: $$W(g^{(\nu+1)}) = \tfrac{1}{2}\int_a^b \left(g^{(\nu)}(y) + \Upsilon w^{(\nu)}(y)\right)dy \int_a^b L(y,t)\left[g^{(\nu)}(t) + \Upsilon w^{(\nu)}(t)\right]dt + \int_a^b v(y)\left\{g^{(\nu)}(y) + \Upsilon w^{(\nu)}(y)\right\}dy.$$

Collecting terms in powers of Υ:

$$W(g^{(\nu+1)}) = W(g^{(\nu)}) + \Upsilon\left\{\tfrac{1}{2}\int_a^b w^{(\nu)}(y)dy\int_a^b L(y,t)\, g^{(\nu)}(t)dt + \int_a^b v(y)\, w^{(\nu)}(y)dy\right\} + \Upsilon^2 \int_a^b w^{(\nu)}(y)dy\int_a^b L(y,t)\, w^{(\nu)}(t)dt;$$

or, using an obviously convenient notation:

$$W(g^{(\nu+1)}) = W(g^{(\nu)}) + \Upsilon W_1(g^{(\nu)}, w^{(\nu)}) + \Upsilon^2 W_2(w^{(\nu)}).$$

Then, completing the square in Υ:

$$W(g^{(\nu+1)}) = W(g^{(\nu)}) + W_2(w^{(\nu)})\left\{\left(\Upsilon + \tfrac{1}{2}\frac{W_1(g^{(\nu)}, w^{(\nu)})}{W_2(w^{(\nu)})}\right)^2 - \tfrac{1}{4}\left(\frac{W_1(g^{(\nu)}, w^{(\nu)})}{W_2(w^{(\nu)})}\right)^2\right\}$$

The largest decrease in $W(g)$ occurs when

$$\Upsilon = -\tfrac{1}{2}\,\frac{W_1(g^{(\nu)}, w^{(\nu)})}{W_2(w^{(\nu)})}.$$

In agreement with the result obtained in linear algebraic systems where the correction vector $W^{(v)}$ is taken along the gradient of a functional W and turns out to be the residue, the function $w^{(v)}(y)$ is taken as:

$$w^{(\nu)}(y) = \int_a^b L(y,t)\, g^{(\nu)}(t)dt + v(y).$$

Hence, the next approximation is:

$$g^{(\nu+1)}(y) = g^{(\nu)}(y) + \Upsilon w^{(\nu)}(y)$$

$$= g^{(\nu)}(y) - \frac{w^{(\nu)}(y)\int_a^b \left(w^{(\nu)}(y)\right)^2 dy}{\int_a^b\int_a^b L(y,t)\, w^{(\nu)}(y)\, w^{(\nu)}(t)\,dy\,dt}.$$

Furthermore the decrease in the functional W is:

$$W(g^{(\nu)}) - W(g^{(\nu+1)}) = \frac{\left(\int_a^b w^{(\nu)}(y)\,dy\right)^2}{2\int_a^b w^{(\nu)}(y)\,dy \int_a^b L(y,t)\,w^{(\nu)}(t)\,dt}$$

Hence the sequence $\{W(g^{(\nu)})\}$ is monotonic decreasing and it has a finite lower bound. Therefore the sequence converges to a unique limit. It follows that

$$\lim_{\nu\to\infty} \left(W(g^{(\nu)}) - W(g^{(\nu+1)}) \right) = 0$$

and the denominator satisfies (II) so that the equation yields the result

$$\lim_{\nu\to\infty} \int_a^b (w^{(\nu)}(y))^2\,dy = 0$$

The sequence $\{w^{(\nu)}(y)\}$ converges to the null function.

Hence: $$\int_a^b L(y,t)\left[g^{(\nu)}(t) - g^{(\nu+1)}(t)\right] dt = 0$$

But, by (i) and by the Schwartz inequality:

$$\left(\int_a^b (g^{(\nu)}(y) - g^{(\nu+1)}(y))^2 dy\right)\left[\int_a^b \left(\int_a^b L(y,t)(g^{(\nu)}(t) - g^{(\nu+p)}(t)\right)^2 dy\right]$$

$$\geq \left(\int_a^b (g^{(\nu)}(y) - g^{(\nu+p)}(y))\,dy \int_a^b L(y,t)(g^{(\nu)}(t) - g^{(\nu+p)}(t))\,dt\right)^2$$

$$\geq m^2 \left(\int_a^b [g^{(\nu)}(y) - g^{(\nu+p)}(y)]^2 dy\right)^2$$

Hence:

$$\left|\int_a^b L(y,t)(g^{(\nu)}(t) - g^{(\nu+p)}(t))\,dt\right| \geq m \left|\int_a^b (g^{(\nu)}(y) - g^{(\nu+p)}(y))\,dy\right|$$

Therefore the sequence, $g^{(\nu)}(y)$, converges to a limit function g(y).

Now if the transformation is assumed to be bounded i.e. satisfies:

$$\left| \int_a^b L(y,t)\, g(t)\, dt \right| \leq K |g(t)|;$$

Then the "triangle" inequality gives:

$$\left| \int_a^b L(y,t)\, g(t)\, dt + v(y) \right| \leq \left| \int_a^b L(y,t)\, g^{(\nu)}(t) + v(y) \right|$$
$$+ \left| \int_a^b L(y,t) \left(g(t) dt - g^{(\nu)}(t) \right) dt \right|$$

and, in the limit, as $v \to \infty$, it is seen that the limit function g(t) satisfies the given equation. for

$$\left| \int_a^b L(y,t)\, g(t)\, dt + v(y) \right| \leq |w^{(\nu)}(y)| + K |g(t) - g^{(\nu)}(t)|.$$

5. A Method of Back Substitution:

The method of back substitution can be applied to the study of integral equations of the first kind:

$$f(x) = \int_a^b L(x,y)\, g(y)\, dy$$

where L (x,y) satisfies the condition of the previous section. An iteration algorithm,

$$g^{(\nu+1)}(y) = g^{(\nu)}(y) + \gamma \left(\int_a^b L(y,x)\, g^{(\nu)}(x)\, dx - f(y) \right),$$

can be easily obtained, for

$$\int_a^b L(x,y)\left[(g(y) - g^{(\nu)}(y)) + g^{(\nu)}(y) \right] dy = f(x).$$

Hence:

$$g(y) - g^{(\nu)}(y) = - \int P(y,x)\, dx$$

where the integral operator $P(y,x)\phi(x)\,dx$ is the operator (if it exists) that has the property

$$\int_a^b P(y,x)\, dx \left(\int_a^b L(x,y)\, p(y)\, dy \right) = p(y)$$

for a p(y) in $L^2(a,b)$. In general, such an operator does not exist; and in every case of interest is not known; for this is equivalent to having solved the problem. Then replace the operator by a real scalar operator, γ, and $g^{(\nu+1)}(y)$ is definedby

$$g^{(\nu+1)}(y) = g^{(\nu)}(y) + \gamma \left[\int_a^b L(y,x)\, g^{(\nu)}(x)\, dx - f(y) \right].$$

This equation is analogous to that obtained in the method of steepest descent; however, in this case γ, being a real constant, does not vary from one iteration to the next. This reduces the rate of convergence but lessens considerably the number of calculations involved. Furthermore, with the kernel satisfying conditions I and II a γ can always be chosen such that the iteration method converges, i.e.

$$\lim_{\nu \to \infty} \left(g^{(\nu+1)}(y) - g^{(\nu)}(y) \right) = 0$$

for almost all y on (a,b). To prove this select

$$g^{(0)}(y) = -\gamma f(y).$$

Then, using an obvious operator formalism,

$$g^{(1)}(y) = \gamma(E+\gamma L) f(y) - \gamma f(y).$$

Hence:

$$g^{(\nu)}(y) = \gamma(E+\gamma L)^{\nu} f(y) - (E+\gamma L)^{\nu-1} f(y) \ldots - \gamma f(y),$$

so that

$$g^{(\nu+1)}(y) - g^{(\nu)}(y) = \gamma(E+\gamma L)^{\nu+1} f(y).$$

Then it is necessary and sufficient for convergence of the iteration process that

$$\lim_{\nu\to\infty} (E+\gamma L)^{\nu+1} = [0]$$

where $[0]$ is the null operator.

Since all of the functions $g^{(\nu)}(y)$, $\nu = 0,1,2,\ldots$ are L^2 (a,b), it is sufficient to consider a function $\phi_i(y)$ satisfying

$$\phi_i(x) = \lambda_i \int_a^b L(x,y)\,\phi_i(y)\,dy.$$

Hence $\lim_{\nu\to\infty} (E+\gamma L)^{\nu}\ \phi_i(y) = \lim (1+\frac{\gamma}{\lambda_i})^{\nu}\ \phi_i(y)$

Since both λ_i and γ are real it is necessary and sufficient that γ satisfy:

$$0 < -\gamma/\lambda_i < 1.$$

Furthermore from the nature of L(x,y) it follows that, since all λ_i are positive, γ must be negative. Let the λ_i be so ordered that

$$\lambda_1 \le \lambda_2 \le \lambda_3 \quad \ldots \quad \lambda_n \le \lambda_{n+1}.$$

Then, convergence is insured if $0 < -\gamma < \lambda_1$. The conditions on L(x,y) insure that $\lambda_1 > 0$.

An obvious estimate for $\lambda_{min.} = \lambda_1$ is obtained by considering the defining equation

$$\phi_1(x) = \lambda_1 \int_a^b L(x,y)\,\phi_1(y)\,dy.$$

Then:

$$\int_a^b \phi_i(x)^2 dx = \lambda_1 \int_a^b \phi_1(x)\,dx \int_a^b L(x,y)\,\phi_1(y)\,dy.$$

By Schwartz' inequality:

$$\int_a^b \phi_i(x)^2 dx \leq \lambda_1 \left(\int_a^b\int_a^b |L(x;y)|^2 dy\,dx\right)^{1/2} \int_a^b \phi_1^2(x)\,dx$$

or, using an accepted notation:

$$\lambda_1 \geq \frac{1}{\|L(x,y)\|}.$$

Hence a possible γ is

$$-\frac{1}{2\|L(x,y)\|}.$$

Furthermore, even if L(x,y) is not "closed" but is "definite", the above estimate will insure convergence.

6. Solution by Quadrature:

A rather obvious approach is to reduce the integral equation to a set of linear algebraic equations, whose solution will presumably give an approximation to the function g(y). This may be done in any of several ways.

Firstly, the integral may be approximated over some set of values $\{y_i\}_{i=1}^{i=n}$ by a suitable quadrature formulated. As an example, use may be made of the Gauss quadrature formula. The range (a,b) is transformed to (-1, 1) and the integral is approximated by:

$$\int_{-1}^{1} K(x,y)\, g(y)\, dy = \sum_{i=1}^{n} \alpha_i^{(n)} K(x, y_i)\, g(y_i)$$

where the values $\alpha_i^{(n)}$ are real constants which depend on the number of points chosen. The set of values $\{y_i\}_{i=1}^{i=n}$ are the n solutions of $P_n(y) = 0$, where $P_n(y)$ is the n^{th} Legendre polynomial. These zeroes are real, distinct, and lie in the range (-1,1). They have been tabulated for fairly large values of n. By this double choice of weight factors and ordinates the integral, evaluated by n points, is exact if the functions integrated are polynomials of degree 2n - 1 or less. For small n this is a decided improvement over the Lagrangian interpolation polynomial. The given function f(x) is then evaluated at the set of points $\{x_i\}_{i=1}^{i=n}$ with each $x_i = y_i$ $i=1,2,\ldots n$ and the system of equations:

$$f(x_j) = \sum_{i=1}^{n} \alpha_i^{(n)} K(x_j, y_i)\, g(y_i) \qquad j = 1, 2, \ldots, n.$$

is solved for $g(y_i)$ $(i = 1,2,\ldots,n)$ the answer being obtained as a set of values $\{g(y_i)\}_{i=1}^{i=n}$. The integral equation may be used to afford some idea of the error. Once the approximations are introduced, there is little point in debating over uniqueness of solution. Obviously this method may well give a solution to the algebraic porblem where there is none to the integral equation. However, if no relation of the form

$$\sum_{i=1}^{n} P_i K(x_j, y_i) = 0 \quad (\text{all } x_j)$$

is satisfied, this method should give a reasonable approximation commensurate with the number of points used. In actuality if the ranges in x and y are different, a set of linear equations with a singular matrix will be obtained providing the intervals in x and y are different.

In addition to the method mentioned any of the iterative schemes for solving simultaneous equations may be used to advantage.

However the important factor of solution stability must be kept in mind. The reduction to a system of algebraic equations involves one approximate quadrature and the remaining calculations are algebraic. An iterative scheme like the steepest descent or back-substitution methods involves repeated integrations each of which is approximate but essentially no solution of a set of linear equations.

It is a matter of extreme complexity to make a precise "error" analysis when an integral equation is solved by numerical methods; however, the need for such an analysis is pressing. Of course once a solution is obtained the error is immediately calculable from the equation itself. An a priori estimation of error, i.e. of the needed fineness of the integration scheme, or of the degree of approximating polynomials, however, is usually a problem of the same order of magnitude as the solution of the given equation. Since iteration methods tend to forestall the unlimited accrueing of round-off and truncation errors which occur in each calculation, it appears that an iterative method would show more stability.

The functions can be maintained as tables of functional values for given ordinates or as tables of coefficients of interpolation polynomials, in which case presumably the quadratures can be performed exactly except for round-off error. Certainly, the case of the kernel, a function two variables, if 100 points were required, 10^4 values need be stored, whereas an interpolation polynomial of 99 degree would require as much storage for its co-efficients.

7. The Method of Fourier Integrals

We wish to obtain a numerical solution to an integral equation of the type

$$f(x) = \int_{-\infty}^{\infty} K(x,y)\, g(y)\, dy.$$

This method will be limited to those particular cases where k(x,y) is of the type k(x-y), k(x+y) or k(xy). By a simple variable transformation each of these can be transformed into k(x-y). So we wish to solve

$$f(x) = \int_{-\infty}^{\infty} K(x-y)\, g(y)\, dy.$$

Formally we can obtain a solution in the following way.
Formally:

$$F(u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{iux} dx$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{iux} dx \int_{-\infty}^{\infty} K(x-y)\, g(y)\, dy$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} g(y)\, dy \int_{-\infty}^{\infty} K(x-y) e^{iux} dx.$$

let x = t + y and obtain

$$F(u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} g(y)\, dy \int_{-\infty}^{\infty} K(t) e^{iu(t+y)} dt$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} g(y)\, dy\, e^{iyu} \int_{-\infty}^{\infty} K(t) e^{iut} dt.$$

Therefore:

$$F(u) = \sqrt{2\pi}\, G(u)\, K(u)$$

where G(u) and K(u) are the Fourier transforms respectively, of g(y) and k(t).

Finally we obtain:

$$G(u) = \frac{1}{\sqrt{2\pi}} \frac{F(u)}{K(u)}$$

and our desired result is

$$g(y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{F(u)}{K(u)} e^{-iyu} du$$

The result is purely formal. The convergence of the improper integrals and the validity of interchange of integration orders must be investigated. The following theorem, however, will be of use in many numerical problems. It is:

Theorem: If: (i) $f(x) \in L^2(-\infty,\infty)$

(ii) $K(x) \in L^2(-\infty,\infty)$

Then if and only if

(iii) $\frac{F(u)}{K(u)} \in L^2(-\infty,\infty)$

is $g(y) \in L^2(-\infty,\infty)$.

We are specifically interested in a numerical solution when the given function f(x) has certain properties:

(i) It is defined, either through lack of complete information or for convenience, at but a finite number of points.

(ii) These points, except perhaps for an assumed constant value at ∞ , are distributed over a finite portion of the real axis.

(iii) It will usually be the case that the values tabulated over the set of points $\{x_i\}$ will not be $f(x_i)$ but $f(x_i) + \epsilon(x_i)$ or even more likely $f(x_i) + \epsilon(x)$, where $\epsilon(x)$ is an error function. That is, the simplest case is where the error at any point is functionally defined by

$$\epsilon(x_j) = \epsilon(j)$$

Usually, through having interpolated by one means of another to obtain each $f(x_i)$, the error at any point x_j, will depend on all points, i.e. $\epsilon = \epsilon(x_j)$.

Naturally we are going to assume that a unique solution to the equation exists. And, due to the incompleteness with which f(x) is given, we must expect to get at best an approximation to g(y).

<u>The Numerical Solution</u>

(a) $$F(u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{iux} dx$$

$$\sqrt{2\pi} F(u) = \int_{-x_0}^{x_0} f(x) e^{iux} dx + I(x_0, u)$$

The cut-off at $x = \pm x_o$ in the integration interval and the subsequent dropping of $I(x_o,u)$ is equivalent to folding the transform $F(u)$ against the transform of the unit step function:

$$m(x) = \begin{cases} 0 & , x > x_0 \\ 1 & -x_0 \le x \le x_0 \\ 0 & x < -x_0 \end{cases}$$

i.e., $\int_{-\infty}^{\infty} F(t) \sin \frac{x_0}{x_0} (u-t) dt$. The precise effect this has on the solution is difficult to estimate. Hence we have here a first limitation. It is necessary that a finite x_o exist such that $I(x_o,u)$ will be sufficiently small. Since $F(u)$ exists this would appear trivial, but if $f(x)$ is a tabulated function or an experimentally measured quantity, knowledge out to x_o, at least, is required. It is possible then then that the range of definition of $f(x)$ may be so limited that $I(x_o,u)$ will be of such a size as to a render a good solution impossible. This difficulty will be discussed later. For the moment we will assume $I(x_o,u)$ to be sufficiently small.

The evaluation of the transforms requires that there be available tables of sines and cosines or that they may be conveniently manufactured in the machine during the process of solution. We assume sets of trigonometric functions will be available.

Further, on the x scale, we assume, since this will usually be most convenient, the values of $f(x)$ to be spaced equidistantly; then let

$$e^{ixu} \rightarrow e^{2\pi i h w} \quad ; \; h = 0, 1, 2, \ldots h_{max}$$

and let $x_0 \rightarrow h_{max}$. Then $h = \frac{h_{max} x}{x_0}$. We have:

$$\sqrt{2\pi}\, F(u) = \int_{-h_{max}}^{h_{max}} [f(x)]_{x\rightarrow h} e^{2\pi i h w} \frac{x_0}{h_{max}} dh + I(h_{max}, w)$$

We can always make dh = 1. Using, in effect the Euler-MacLaurin summation formula:

$$\sqrt{2\pi}\, F(u)_{u\rightarrow w} \simeq \frac{x_0}{h_{max}} \sum_{-h_{max}}^{h_{max}} [f(x)]_{x\rightarrow h} e^{2\pi i h w} .$$

Similarly, for the kernel,

$$\sqrt{2\pi}\, K(u)_{u\rightarrow w} \simeq \frac{x_0}{h_{max}} \sum_{-h_{max}}^{h_{max}} [K(x)]_{x\rightarrow h} e^{2\pi i h w}$$

Form the quotient:

$$\frac{\sum_{-h_{max}}^{h_{max}} [f(x)]_{x\to h} e^{2\pi i h w}}{\sum_{-h_{max}}^{h_{max}} [k(x)]_{x\to h} e^{2\pi i h w}} = B_j : \; j = 0, 1, 2, \ldots \frac{w_{max}}{\Delta w}$$

which is to represent $\frac{F(u)}{K(u)}$ at $\frac{w_{max}}{\Delta w}$ values of w corresponding to u.

That means that for $0 \leq w \leq w_{max}$, u runs over $0 \leq u \leq \frac{2\pi h_{max}}{X_0} W_{max}$.

Next we form:

$$g(y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} G(u) e^{-iyu} du = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{F(u)}{K(u)} e^{-iyu} du$$

$$2\pi g(y) = \int_{-u_{max}}^{u_{max}} \frac{F(u)}{K(u)} e^{-iyu} du \mp J(u_{max}, y).$$

The Euler-MacLaurin integration scheme gives:

$$2\pi g(y)_{y\to z} = \frac{u_{max}}{m_{max}} \sum_{m=-u_{max}}^{m=+u_{max}} \left[\frac{F(u)}{K(u)}\right]_{u\to m} e^{-2\pi i m z} \Delta m$$

and again let $\Delta m = 1$. The variable ranges are as follows:

(i) x: 0 to x_0

(ii) h: 0 to $H(= h_{max})$; $\Delta h = 1$

(iii) w: 0 to $W(= w_{max})$ (say 0.250 of a cycle)

hence from H values of the integrand, we get

$\frac{W}{\Delta w}$ values of F(u).

(iv) u: 0 to $\frac{2\pi HW}{X_0}$

of the $\frac{W}{\Delta w}$ values of $\left[\frac{F(u)}{K(u)}\right]_{u\to w}$ we may use all, or an equidistant subset in evaluating g(y).

Hence while (iv) holds,

(v) m: 0 to M in M jumps; $\Delta m = 1$.

and $\frac{W}{\Delta w} = Ml$; l an integer.

(vi) z: 0 to $Z(= z_{max})$

hence from M values of $\left[\frac{F(u)}{K(u)}\right]_{u \to w}$ we obtain

$Z/\Delta z$ values of $g(y)_{y \to z}$

for which:

(vii) y: 0 to $\frac{2\pi M z}{2\pi H W} X_0 = \frac{MZ}{HW} X_0$.

Having obtained g(y), the problem is essentially solved.

More may be said about the actual evaluation of the transforms. We can always choose our zero on the x axis so as to maintain the following symmetry:

$$F(u) = \frac{1}{\sqrt{2\pi}} \int_0^{X_0} \left\{ [f(x) + f(-x)] \cos ux + i [f(x) - f(-x)] \sin ux \right\} dx$$

and likewise for K(u). Both F(u) and K(u) are evaluated over the same sets of points. And

$$G(u) = \frac{1}{\sqrt{2\pi}} \frac{F_{cos}(u) + i F_{sin}(u)}{K_{cos}(u) + i K_{sin}(u)} = \frac{1}{\sqrt{2\pi}} \frac{(F_c + i F_s)(K_c - i K_s)}{K_c^2(u) + K_s^2(u)}.$$

Since we integrate over $(-\infty, \infty)$ all odd terms vanish. Hence we must furnish:

(i) $K_c^2(u) + K_s^2(u)$

(ii) $F_c K_c$

(iii) $-F_s K_s$

and

g(y) over the range 0 to y_{max} will be the (cosine + sine) transform of G(u), while being the (cosine - sine) transform of G(u) over $-y_{max}$ to 0.

Cut-off errors: In general, it is very difficult to assess the error in the solution of this problem. The way they arise is at once evident though their accumulative effect on the solution is difficult to determine. The errors are the usual ones in a numerical calculation. We may list them here. They are:

(i) The function f(x), either through expediency, or in fact, is incompletely defined an, as mentioned before, will never, except possibly for a constant value, be given over the entire range $(-\infty, \infty)$.

(ii) In addition, each value, $f(x_i)$ will be contaminated by errors. These will be, at least, truncation errors.

(iii) The infinite integrals must always, unless they can be evaluated exactly by some analytic procedure, be cut off at some finite value. This is equivalent, as mentioned before, to the folding of the precise transform against the step function $\frac{\sin ux_0}{ux_0}$ if x_0 is the cut-off value. This means that the frequency spectrum of f(x) is warped by this factor. What is more important, though, the warping cannot be removed form g(y) by any unfolding.

However, we can make some assumptions about f(x) for large X and develop an error formula for the cut-off. We observe that we have two cut-off integrals to evaluate:

(a) $$\int_{x_0}^{\infty} f(-x)\, e^{-iux}\, dx$$

(b) $$\int_{x_0}^{\infty} f(x)\, e^{iux}\, dx$$

As $x_0 \to \infty$, these errors $\to 0$; and the high frequency components of F(u) should be largely transformed at any reasonably large x_0 and similarly for the cut-off error in the evaluation of K(u)

The integration scheme

The evaluation of F(u) and G(u) is to be done numerically and in those cases where it cannot easily be evaluated analytically K(u) will be also so treated.

Since we are evaluating special types of integrals, i.e. Fourier transforms, it is best that we use an integration scheme which allows us to make most use of the duplicative properties of sines and cosines.

To make best use of the computer's limited storage capacity we must use a process that forms the integration components from a minimum storage list of sines and cosines. For the moment we will stand pat on the Euler-MacLaurin integration scheme using equidistantly spaced points. It is:

$$\int_1^n f(t)dt = \frac{1}{2}[f(1)+f(n)] + \sum_{n=2}^{n-1} f(n) - \int_1^n f'(t) S_1(t)dt$$

where

$$S_1(t) = -\frac{1}{2}\sum_{n=1}^{\infty} \frac{\sin 2\pi nt}{2n\pi}$$

We are to evaluate an integral of the form:

$$\int_0^{x_0} f(x)e^{iux}dx = \frac{W_0}{H}\int_0^H f(H)e^{2\pi ihw}dh$$

where $H = X_0 = h_{max}$

The sum formula yields:

$$\int_0^H f(h)e^{2\pi ihw}dh = \frac{1}{2}[f(0) + f(H)e^{2\pi iHw}] + \sum_{n=1}^{H-1} f(n)e^{2\pi inw} - \int_0^H \frac{d}{dh}(f(h)e^{2\pi ihw})S_1(h)dh.$$

An analysis of the error term is desirable for it develops certain facts about the relation of the h and w intervals. Naturally as the number of intervals $\rightarrow \infty$, the function S(h) $\rightarrow$ 0. The functional relationship between the error and the interval length resides, of course in S(h). We will assume that $|f(h)| < M$, M a finite constant, over(0,H). This is not unreasonable. We consider the error for W = 0:

$$E(0,H) = M\int_0^H \left(\sum_1^{\infty} \frac{1}{2n\pi}\sin 2n\pi h\right)dh$$

$$= M\int_0^H (h - \tfrac{1}{2} - [h])dh = 0$$

([h] is the least integer in (0,1), (1,2),...etc)
For W = m, m = integer

$$E(M,H) = \sum_{n=1}^{\infty}\left(\frac{1}{2n\pi}\right)\frac{M}{2\pi im}\int_0^H e^{2\pi imh}\sin 2\pi nh\, dh = \frac{MH}{4m^2\pi}$$

For the error to be zero, except at $W = \infty$, it is necessary that $e^{2\pi i h w}$ be orthogonal to

$$\sum_{n=1}^{\infty} \frac{\sin 2n\pi h}{2n\pi}$$ over (0,H). This it cannot be.

Hence it would appear that, since $W = 0$ is a value we use, we should though the peak about $W = 1$ is narrow, limit W to lie certainly in the range $0 \leq W \leq C < 1$, to minimize the error. This of course is a consequence of the fact that we are fortunate to be dealing with periodic functions, $e^{2\pi i h w}$, super-imposed upon, as we have assumed, a function containing practically no frequencies above some finite cut-off point. This concludes the analytic discussion of the errors exclusive of those occurring from round-off, truncation, or in the function f(x) itself. These latter can probably be treated statistically and this may actually be done later.

Convergence Factors

Before going on to the more numerical treatment we must mention some important considerations concerning convergence. We assume that F(u) and K(u) exist and are finite. Hence F(u) must have a higher order zero at infinity than K(u) to insure convergence of

$$g(y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{F(u)}{K(u)} e^{-iyu} du$$

But it is precisely those features which we lack in any approximate solution which make this extremely difficult to achieve. Every error that occurs in the procedure has the effect of adding into F(u), Fourier components which extend the range in u on which F(u) maintains appreciable value. It is only through a sacrifice in "resolution" of g(y) that we will be able to obtain a convergent numerical solution. Nevertheless, admitting the necessity, and even justice, of such a resolution depletion we can obtain solutions which are quite good. The extent of this "smudging out" of g(y) will depend, then, on the nature of g(y), the extent of definition of f(x), and, of course, the accuracy we maintain in the numerical computations. A detailed example will be shown later where the function g(y) is a sum of δ- functions. Even though its transform does not converge we are able to obtain a partially resolved solution; indeed in this worst possible case we do obtain a resolution as fine as desirable by merely extracting an increasing number of values of f(x) and refining our integration scheme, accordingly. Hence to obtain a g(y), we must usually expect to be forced to introduce a convergence factor. A convenient one, since up to constant amplitude and half-width factors it transforms into itself, is e^{-ax^2}. Another is $\frac{\sin^2 ax}{(ax)^2}$ and still a third is $x^2 e^{-ax^2}$. As in the formation of the first transforms F(u), K(u), this "folding function" or convergence factor gives us a solution at the expense of resolution. For, since we cannot calculate

$$g(y) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{F(u)}{K(u)} e^{-iyu} du.$$

We introduce the factor T(u) and get:

$$g^*(y) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{F(u)}{K(u)} T(u) e^{-iyu} du$$

$$= \int_{-\infty}^{\infty} g(y-s)\, t(s)\, ds.$$

In short, our solution is "folded against" the transform of T(u). The step function yields a $g^*(y)$ which has oscillations of decreasing amplitudes and obscures much of the result. The function e^{-au^2} transforms into $e^{-\frac{1}{4a}s^2}$, up to an amplitude factor. This means that the greater the value is u at which $\frac{F(u)}{K(u)}$ remains well behaved, the narrower will be the spread by which each point will be weighted in the $g^*(y)$. This is almost poetic in its justice. For in just this way does the incomplete definition of f(x) affect the result.

A specific example

Consider the Laplace integral equation

$$f(\xi) = \int_0^{\infty} e^{-\xi\eta} g(\eta)\, d\eta$$

The transformation

$$\eta = e^{-y}, \quad \xi = e^{x}$$

yields

$$e^x f(e^x) = \int_{-\infty}^{\infty} e^{-e^{(x-y)}} e^{x-y} g(e^{-y})\, dy$$

or

$$e^x \tilde{f}(x) = \int_{-\infty}^{\infty} e^{-e^{(x-y)}} e^{x-y} \tilde{g}(y)\, dy.$$

We form:

$$F(u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^x \tilde{f}(x) e^{ixu} dx$$

$$K(u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-e^t} e^t e^{iut} dt$$

and their quotient to give:

$$\tilde{g}(y) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{F(u)}{K(u)} e^{-iyu} du.$$

With this kernel, K(u) can be evaluated analytically and is easily seen to be the Euler integral for $\frac{1}{\sqrt{\pi}} \Gamma(1 + iu)$. Likewise a necessary and sufficient condition for the behavior of f(x) for large x(i.e. x_0) can be stated. With respect to the two cut off integrals it is as follows: The cut off integrals are:

(a) $$\int_{x_0}^{\infty} e^{-x}\tilde{f}(-x) e^{-iux} dx$$

(b) $$\int_{x_0}^{\infty} e^{x}\tilde{f}(x) e^{iux} dx.$$

Then for convergence, it is both necessary and sufficient that, for (a), for large x, f(-x), be $O(e^{nx})$, $n < 1$ or f(x) be $O(e^{-nx})$, $n < 1$.; and that for (b), for large x, f(x) be $O(e^{nx})$; $n < -1$. For both these cases an error formula can be simply obtained by a single integration by parts. They are:

(a) $$\left|\int_{x_0}^{\infty} e^{-x} O(e^{nx}) e^{-iux} dx\right| \le \left|\frac{1}{n-1} O(e^{(n-1)x_0})\right| + \frac{|u|}{|n-1|}\left|\int_{x_0}^{\infty} O(e^{(n-1)x_0}) e^{-iux} dx\right|$$

$$\left|\int_{x_0}^{\infty} e^{-x} O(e^{nx}) e^{-iux} dx\right| \le \frac{O(e^{(n-1)x_0})}{|n-1| - |u|}.$$

and (b) yields similarly:

$$\left|\int_{x_0}^{\infty} e^{x} f(x) e^{iux} dx\right| \le \frac{O e^{(n+1)x_0}}{|n+1| - |u|}$$

A more specific result holds for the cut-off in forming K(u). We observe that these two formulas yield an exact expression for the error in the two cut-off's. They are:

$$\int_{x_0}^{\infty} e^{-e^t} e^t e^{itu} dt = -\Gamma_{\ln x_0}(1+iu) + \Gamma(1+iu)$$

and $$\int_{-\infty}^{-x_0} e^{-e^t} e^t e^{itu} dt = -\Gamma_{\ln \frac{1}{x_0}}(1+iu)$$

where $\Gamma_j(x)$ referes to the incomplete gamma function of x with integration limit j.

Having formed the quotient $\frac{F(u)}{\Gamma(1+iu)}$ we observe that:

$$|\Gamma(1+iu)| = \frac{\sqrt{\pi}}{|u|} \frac{1}{|\sinh \pi u|}$$

so that K(u) diminishes quite rapidly. Assuming that a convergence factor is necessary let us introduce e^{-au^2} as one. Our choice of "a" is governed by the behaviour of $\frac{F(u)}{K(u)}$ as u increases. The necessary existence of undamped Fourier terms in F(u) implies that for some value of u, the quotient will begin to grow without bound. We select an "a" such that $\frac{F(u)}{K(u)} e^{-au^2}$ will be essentially zero for all u beyond this critical value. Hence we have:

$$\tilde{g}^*(y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{F(u)}{\Gamma(1+iu)} e^{-au^2} e^{-iuy} du$$

$$= \frac{1}{\sqrt{2a}} \int_{-\infty}^{\infty} \tilde{g}(y-t)\, e^{-(t/2a)^2} dt$$

and we see at once the way in which a small "a" (i.e. a large critical "u") yields a good $\tilde{g}^*(y)$. Here then is the crux of the resolution problem. It is all bound up in the constant "a". The size of "a" will then depend upon just how well we can damp out the "noise" in the evaluation of F(u). That it can be made to be as near zero as we wish is a consequence of the uniqueness of the solution.

8. The Fredholm Integral Equation of the Second Kind:

1. The Existence Theorem:

The equation to be considered is:

$$w(x) = g(x) + \lambda \int_a^b K(x,t) w(t) dt.$$

The functions w(x), f(x), and K(x,t) are assumed to be real functions of the real variables x,y. As in the preceeding section, if K(x,t) is not symmetrical it may be made to yield a symmetric kernel. The above equation may be written in symbolic form:

$$w(x) = g(x) + \lambda T w(x).$$

or

$$g(x) = (E - \lambda T) w(x) \qquad (*)$$

where E is the identity operator and T is the integral operator. Obviously the equation is solved if $(E - \lambda T)^{-1}$ can be obtained. In general, it will be shown that such an inverse exists, subject to certain conditions, as an infinite series in powers of T(i.e. as sums of iterated integrals).

Left multiplication of both sides of (*) by $(E - \lambda T)'$ (the transpose) yeilds

$$u(x) = (E - \lambda L) g(x) \qquad (\dagger)$$

where L is a symmetric integral operator and hence has only real eigenvalues. The equation to be considered then is either (*) or (†). For these equations the following existence theorem is fundamental:

If the known functions g(x) and K(x,t) are integrable squared over (a,b), then is w(x) integrable squared over (a,b). Furthermore if λ is such that the homogeneous equation

$$\phi(x) = \lambda T \phi(x)$$

has only the solution $\phi(x) = 0$, then there exists a unique solution w(x). If, however, the homogeneous equation has m independant solutions then there exists a w(x) if and only if g(x) is orthogonal to each of the m solutions of

$$\psi(x) = \lambda T' \psi(x)$$

The solution w(x) is not unique for there can be adjoined to it any linear combination of the m solutions of the homogeneous equation.

2. The Simple Iteration Procedure

We form, from:

$$u(x) = f(x) + \lambda \int_a^b L(x,t) u(t) dt$$

the iteration procedure:

$$u^{(n)}(x) = f(x) + \lambda T u^{(n-1)}(x).$$

then

$$u_r^{(n)}(x) = f(x) + \lambda T f(x) + \lambda^2 T^2 f(x) + \ldots + \lambda^{n-1} T^{n-1} f(x) + R_n$$

where

$$R_n = \lambda^n T^n u_0(x).$$

let $|u_0(x)| \leq U$, then

$$\|R_n\| \leq |\lambda^n| |U| \|T^n\|$$

a necessary and sufficient condition for convergence is that $\lim_{n\to\infty} \| R_n \| = 0$ so that:

$$\lim_{n\to\infty} |\lambda^n| \|T^n\| \to 0 ;$$

a sufficient condition is, of course,

$$|\lambda| \|T\| < 1$$

or:

$$|\lambda| \|L\| < 1$$

We define the morm of T by the following equations:

$$\|L\| = \left(\int_a^b\int_a^b |L(x,y)|^2 dx dy \right)^{1/2}$$

and

$$\int_a^b L(x,y) u(y) dy = T u(y)$$

We define:

$$\frac{\|T u(x)\|}{\|u(x)\|} \leq \|T\|$$ i.e. the least number for which the inequality is true. It follows that

$$\|Tu\|^2 \leq \|L\|^2 \|u(x)\|^2$$

from the Schwartz inequality. Hence:

$$\|T\| \leq \|L\|$$

Furthermore

$$\|T^K u(x)\| \leq \|T\| \, \|T^{K-1} u(x)\|.$$

Hence, if

$$S_n = 1 + \sum_{n=1}^{\infty} T^n$$

we have

$$\|S_n\| \leq 1 + \|T\| + \|T^2\| +$$

The solution so obtained is unique. A major problem, then, is how to make use of an iterative process when, say, $\| \lambda L \| \geq 1$.

Let $\| \lambda L(x,y) \|$ be bounded. Then an algebraic procedure can be adjoined to the iteration procedure to provide a solution when $\| \lambda L(x,y) \| \geq 1$. It is always possible to find a set of functions $A_i(x)$, $B_i(x)$ such that:

$$\| L(x,y) - \sum_{i=1}^{N} A_i(x) B_i(y) \| < 1$$

or

$$L(x,y) = \sum_{i=1}^{N} A_i(x) B_i(y) + M(x,y)$$

where N is chosen so that $\| M(x,y) \| < 1$. Hence $M(x,y)$ has an inverse, J, consisting of convergent infinite series of integral operators. Then:

$$u(x) = J f(x) + \sum_{i=1}^{N} (J A_i(x)) \int_a^b B_i(y) u(y) dy,$$

or, using an obvious notation:

$$(B_j(x), u(x)) = (B_j(x), J f(x)) + \sum_{i=1}^{N} (B_j(x), J A_i(x)) (B_i(y), u(y)); \quad j = 1, 2, \cdots, N$$

$$(B_j(x), u(x)) = \int_a^b B_j(x) u(x) dx$$

The solution this set of linear equations for $(B_j(x), u(x))$ when substituted in the equation for u(x) yields the desired solution u(x). This method, in essence, can be regarded as a double iteration scheme, if the algebraic equations are so solved.

9. The Method of Steepest Descent

The equation

$$f(x) = g(x) + \lambda \int_a^b K(x,y) f(y) dy$$

in operator form is

$$(E - \lambda T) f(x) = g(x)$$

and T will be assumed to be a positive operator. Then, it follows that the inequalities

$$0 < m \int_a^b f(x)^2 dx \leq \int_a^b f(x) dx \int_a^b K(x,y) f(y) dy \leq M \int_a^b f(x)^2 dx$$

are satisfied, the first from the positive nature of the operator, the second from the bounded nature of the operator and Schwartz's inequality.

The functional w(f) is defined by

$$W(f) = 1/2 \int_a^b \{ f(x)^2 - \lambda f(x) \int_a^b K(x,y) f(y) dy \} dx + \int_a^b f(x) g(x) dx$$

Squaring the expression

$$\int_a^b \{ f^2(x) - \lambda f(x) \int_a^b K(x,y) f(y) dy \} dx + \int_a^b f(x) g(x) dx.$$

yields the inequality

$$2 \left(\int_a^b \{ f^2(x) - \lambda f(x) \int_a^b K(x,y) f(y) dy \} dx \right) W(f) + \left(\int_a^b f(x) g(x) dx \right)^2 \geq 0$$

from which follows

$$2m \left(\int_a^b f^2(x) dx \right) W(f) + \int_a^b f^2(x) dx \int_a^b g^2(x) dx \geq 0.$$

Therefore w(f) has a finite lower bound, independant of f(x), given by

$$W(f) \geq - \frac{\int_a^b g^2(x) dx}{m}$$

If $f^{(v)}(x)$ is an approximation, then an improvement, $f^{(v+1)}(x)$ is obtained by considering

$$W(f^{(v+1)}) = W(f^{(v)}(x) + \alpha r^{(v)}(x))$$

where α, a real number, and $r^{(v)}(x)$ are to be determined. Then

$$W(f^{(v+1)}) = W(f^{(v)}) + \alpha \frac{1}{2} \int_a^b \{ r^{(v)}(x) f^{(v)}(x) - \lambda r^{(v)}(x) \int_a^b K(x,y) f^{(v)}(y) dy \} dx$$

$$+ \alpha \int_a^b g(x) r^{(v)}(x) dx + \alpha^2 \int_a^b \{ r^{(v)}(x)^2 - \lambda r^{(v)}(x) \int_a^b K(x,y) r^{(v)}(y) dy \} dx.$$

or, in shorter notation:

$$W(f^{(\nu+1)}) = W(f^{(\nu)}) + \alpha\Theta_1 + \alpha^2\Theta_2$$

Completing the square in α:

$$W(f^{(\nu+1)}) - W(f^{(\nu)}) = \Theta_2\left(\alpha + \tfrac{1}{2}\frac{\Theta_1}{\Theta_2}\right)^2 - \tfrac{1}{4}\left(\frac{\Theta_1}{\Theta_2}\right)^2.$$

w(f) decreases most rapidly when

$$\alpha = -\tfrac{1}{2}\frac{\Theta_1}{\Theta_2}$$

as before, the function $r^{(\nu)}(x)$ is taken as the residue, i.e.

$$r^{(\nu)}(x) = f^{(\nu)}(x) - \lambda\int_a^b K(x,y)\, f^{(\nu)}(y)\,dy - g(x).$$

The next approximation is

$$f^{(\nu+1)}(x) = f^{(\nu)}(x) + \alpha\, r^{(\nu)}(x)$$

and this is the iteration algorithm. The sequence, $\{ w(f^{(\nu)}(x))\}$, being monotonic decreasing, has a finite lower bound and so converges to a unique limit, from which it follows that

$$\lim_{\nu\to\infty} \left[W(f^{(\nu+1)}) - W(f^{(\nu)})\right] = 0$$

which by a previous inequality implies

$$\lim_{\nu\to\infty} \int_a^b r^{(\nu)}(x)^2\,dx = 0$$

from which it follows that $\{ f^{(\nu)}(x)\}$ converges to the solution f(x), almost everywhere.

10. The Method of Back-Substitution

The equation to be considered is:

$$f(x) = g(x) + \lambda \int_a^b K(x,y) f(y) dy$$

The operator in question may be made "definite by reduction to:

$$f(y) = \left[g(y) + \lambda \int_a^b K(x,y) g(x) dx \right] + \lambda^2 \int_a^b K(x,y) dx \int_a^b K(x,y) f(y) dy.$$

or

$$f(y) = u(y) + \lambda^2 \int_a^b L(x,y) f(x) dx.$$

The iteration scheme is:

$$f^{(0)}(y) = -\alpha\, g(y)$$

and

$$f^{(\nu+1)}(y) = f^{(\nu)}(y) + \alpha r^{(\nu)}(y)$$

where

$$r^{(\nu)}(y) = f^{(\nu)}(y) - \lambda \int_a^b L(x,y) f^{(\nu)}(x) dx - g(y)$$

and α is to be chosen so that $\lim_{\nu \to \infty} | r^{(\nu)}(y) | = 0$

Then:

$$r^{(\nu+1)}(y) = f^{(\nu+1)}(y) - \lambda \int_a^b L(x,y) f^{(\nu+1)}(x) dx - g(y)$$

$$= f^{(\nu)}(y) + \alpha r^{(\nu)}(y) - \lambda \int_a^b L(x,y) \left[f^{(\nu)}(x) + \alpha r^{(\nu)}(x) \right] dx - g(y).$$

and

$$r^{(\nu+1)}(y) = r^{(\nu)}(y) + \alpha r^{(\nu)}(y) - \lambda \alpha \int_a^b L(x,y) r^{(\nu)}(x) dx$$

or, in operator notation:

$$r^{(\nu+1)}(y) = (E + \alpha(E - \lambda T)) r^{(\nu)}(y).$$

Since the operators are bounded and commute this becomes:

$$r^{(\nu+1)}(y) = \left[E + \alpha(E - \lambda T) \right]^{\nu} r^{(1)}(y)$$

and α must now be chosen so that $\lim_{\nu \to \infty} r^{(\nu+1)}(y) = 0$. Since the functions considered are $L^2(a,b)$ and T is a bounded "positive" operator, it will be assumed that

$$0 < m \int_a^b f^2(x) dx \le \int_a^b f(x) dx \int_a^b L(x,y) f(y) dy \le M \int_a^b f^2(x) dx$$

for any f(x) in $L^2(a,b)$. From the above it follows that

$$\| r^{(\nu+1)}(y) \| \leq \| E + \alpha(E - \lambda T) \|^{\nu} \, \| r^{(1)}(y) \|$$

Hence $\lim_{\nu \to \infty} \| r^{(\nu+1)}(y) \| = 0$

if and only if

$$\| E + \alpha(E - \lambda T) \| < 1$$

or

$$\int_a^b f^2(x)\,dx + \alpha \int_a^b f^2(x)\,dx - \alpha\lambda \int_a^b f(x)\,dx \int_a^b L(x,y) f(y)\,dy < \int_a^b f^2(x)\,dx$$

Then,

$$| 1 + \alpha - \alpha\lambda m | < 1$$

yields a relation for α. At this point it might be mentioned that the convergence criterion is the same as for Newton's method, with $\alpha = -\gamma$ Hence, subject to this change in sign, the two methods converge or diverge together. This method also fails, then, for $\lambda = \lambda_i$, a characteristic value of the homogeneous equation. Newton's method will, in general, however, converge more rapidly.

11. The Solution by Newton's Method

The equation

$$f(x) = g(x) + \lambda \int_a^b K(x,t)\, f(t)\, dt$$

can be represented in operator form as

$$(E - \lambda T) f(x) = g(x)$$

defined on the interval (a,b), where E is the identity operator and T is a positive integral operator. An operator $(E - \lambda T)^{-1}$ exists and in the case of $\| \lambda K(x,t) \| < 1$ is representable as an infinite series in iterated integrals operating on g(x). Newton's method may be employed to obtain $(E - \lambda T)^{-1}$, in those cases where $\| \lambda K(x,t) \| \geq 1$.

By analogy to the Newton algorithm for obtaining reciprocals of numbers, the equation

$$p^{(v+1)} = p^{(v)} [2E - (E - \lambda T) p^{(v)}]$$

is formed where the operator $p^{(v)}$ is such that

$$\lim_{v \to \infty} p^{(v)} = (E - \lambda T)^{-1}.$$

Considerations of commutativity may be ignored if $p^{(0)} = \gamma$, a scalar. Then $p^{(0)}$, and hence, by induction, $p^{(v)}$ $(v = 1,2,\ldots)$ commute with the integral operator λT. This is quite obviously a sufficient, and even convenient, but not necessary constraint on $p^{(v)}$ $(v = 0, 1,2,\ldots)$. By applying the recurrence relation it is seen that

$$E - (E - \lambda T) p^{(v)} = (E - (E - \lambda T) p^{(0)})^{2^v}$$

Hence

$$\lim_{v \to \infty} p^{(v)} = (E - \lambda T)^{-1}$$

if and only if

$$\lim_{v \to \infty} \| E - (E - \lambda T) p^{(0)} \|^{2^v} = 0$$

A sufficient condition that this be so is that

$$\| E - (E - \lambda T) p^{(0)} \| < 1$$

By a proper choice of $p^{(o)} = \gamma$ this can always be arranged. Hence Newton's method wil always yield a solution by a purely iterative scheme (providing of course the operator T is bounded, which will be assumed at all times). Obviously if $\lambda = \lambda_i$ is a characteristic value of the equation, the inequality above is not satisfied. Likewise such a λ violates a necessary condition, i.e. that all characteristic values of the operator $E - (E - \lambda T) p^{(o)}$ lie outside the unit circle in the complex plane for with a $\lambda = \lambda_i$ the operator has a characteristic value on the unit circle.

12. Solution by Approximate Quadrature

The Fredholm equation

$$f(x) = g(x) + \lambda \int_a^b K(x,y)\, f(y)\, dy$$

may be regarded (and was so treated originally by Fredholm) as the limiting case of n simultaneous linear algebraic equations

$$f(x_i) = g(x_i) + (\lambda/n) \sum_{j=1}^{n} K(x_i, y_j) f(y_j) \qquad i = 1, 2, \ldots, n.$$

as $n \to \infty$. If λ is not a characteristic value of the homogeneous equation, the above, for n large enough, may be assumed to yeeld a solution $f(x_i)$ $i = 1,2,\ldots,n$, which clossely approximates f(x) at least at n points. Since a solution can only be obtained for a finite number of points, in any case, when numerical methods are employed the above appears to be an attractive way to dispose of the Fredholm equation by reducing the transcendental to an algebraic problem.

The above equation represents a solution at "n" distinct points equidistantly spaced. However, it involves the most primitive quadrature method. A linear quadrature scheme which best approximates the integral is desired. One method is to approximate the functions involved by interpolation polynomials, and prevail upon an exact integration to give the algebraic system. Lagrangian or orthogonal polynomials are most commonly employed. Another method involves the use of non equidistantly spaced points and the concept of Gaussian mechanical quadrature. Firstly the interval (a,b) is mapped into (-1,1). The integral is approximated

$$\int_{-1}^{1} K(x,y)\, f(y)\, dy = \sum_{j=1}^{n} \alpha_j^{(n)} K(x, y_j^{(n)})\, f(y_j^{(n)}).$$

with a given fixed n. The $\alpha_j^{(n)}$ are weights which vary with j and each n. The points $y^{(n)}$, $j = 1,2,\ldots n$ are the n real, distinct, zeroes of the nth Legendre polynomial and lie in (-1,1). Tables of the $y_j^{(n)}$ are available for $n \leq 10$ and, in time, will be no doubt available for larger n. Such a quadrature yields an exact evaluation of the integral if the intengrand is a polynomial of degree 2n - 1 by the use of only n points.

The method may be generalized to treat any interval providing the points and weights are chosen to depend upon that polynomial orthogonal over (a,b) with respect to a unit weight function. The more common orthogonal functions may be used providing the proper weight functions are introduced. For example, on (-1,1) the Tchebichef polynomials may be used if the integral is written as:

$$\int_{-1}^{1} \frac{[K(x,y)\, f(y) \sqrt{1-y^2}\,]}{\sqrt{1-y^2}}\, dy = \sum_{i=1}^{n} \beta_i^{(n)} K(x, y_i^{(n)})\, f(y_i^{(n)}) \sqrt{1-y_i^{(n)2}}$$

Here the $\beta_i^{(n)}$ are identical for all i and a given n, though the $y_i^{(n)}$ are not equidistantly spaced being the n zeroes of the Tchebichef polynomial $T_n(y)$. Likewise use may be made of the Laguerre polynomials over $(0,\infty)$ with weight function $e^{-x/2}$ and Hermite polynomials over $(-\infty, \infty)$ with weight $e^{-x^2/2}$.

Bibliography

T. Aheroni, "Antennae, An Introduction to Their Theory", Oxford, 260 pp. (1946).

H. Bateman, "On definite functions," Mess. Math., 37, 91-95 (1908).

H. Bateman, "A formula for the solving function of a certain integral equation of the second kind", Mess. Math., 37, 179-186 (1908).

H. Bateman, "Notes on Integral Equations", Mess. Math., 38, 8-13, 70-75, (1909)

H. Bateman, "The Theory of Integr-l Equations," Proc. Lond. Math. Soc. 4, 90-115 (1907).

H. Bateman, "On the Inversion of a Definite Integral," Proc. Lond. Math. Soc. 4, 461-498, (1907).

P. D. Crout, "An Application of Polynomial Approximation to the Solution of Integral Equations Arising in Physical Problems", Jour. Math. Phys. 19, 34-92 (1940).

J. B. Diaz and Alexander Weinstein, "Schwarz Inequality and Methods of Rayleigh-Ritz and Trefftz", Journ. Math. and Phys., 26, 133-136 (1947).

E. Picard, "Sur un Theoreme general relatif aux equations integrales de fremiere espece et sw quelques problemes de Physique mathematique," Rend. Circ. Math. Palermo, 29, 79-97 (1910).

E. Hellinger and O. Toeplitz, "Integral gleichungen und Gleichungen mit unendlichirelen Unbekannten", Encyi. Math. Wiss. 2 (3), 1335-1597 (1927).

Maurice Janet, "Equations integrales et applications a certains problemes de la physique mathematieque", Mem. Sci. Math., 101-102, 150 pp. (1941).

W. V. Lovitt, "Linear Integral Equations" McGyrais-Hill, (1924).

E. J. Mystrom, "Uber die Praktische Aublosung von Integral gleichungen mit Arrivendungen auf Randwertaufgaben", Acta. Math. 54, 185-204 (1930).

G. Prasad, "On the numerical solution of integral equations," Proc. Edin. Math. Soc. 42, 46-59 (1924).

A. Reiz, "On the Numerical Solution of Certain types of Integral Equations" Arkiv. for Math. 29A, 21 (1943).

Erhard Schmidt, "Zur Theorie der linearen und nicht linearen integral gleichungen," Math. Ann. 63, 64, (1907).

M. J. Soula, "L'Equation Integrale de Premiere Espece a Limites et Les Fonctions Permutables a Limites Fixes," Mem. Sci. Math., 80, 63 pp. (1936).

G. Temple, "The General Theory of Relaxation Methods Applied to Linear Systems" Proc. Roy Soc. Lond. 169A, 476-500 (1939).

E. C. Tichmarsh, "Theory of Fourier Integrals", Oxford (1937).

E. T. Whittaker, "On the Numerical Solution of Integral Equations," Proc. Roy. Soc. Lond. 94A, 367-383 (1918).

Signed: Alan J Perlis
Alan J. Perlis

Approved: Philip Franklin
Philip Franklin

AJP:bno

6345
Engineering Note E-328

leBian
~~Zahnestack~~

Project Whirlwind
Servomechanisms Laboratory
Massachusetts Institute of Technology
Cambridge, Massachusetts

SUBJECT: PULSE TRANSFORMERS AND INTERSTAGE COUPLING IN WHIRLWIND I

To: Systems Group, Storage Tube Group

From: C. A. Rowland

Date: January 31, 1950

Abstract: Pulse transformers are useful for coupling between pulse circuits of a system because they permit a saving in the number and size of the tubes required.

A one-to-one inverter transformer is useful for coupling between adjacent stages, and a step-down step-up arrangement is useful for coupling remote stages. Crystal diodes are useful in coupling circuits for damping, isolating, and clamping purposes; the application of crystal diodes is particularly helpful in point-to-multipoint and multipoint-to-multipoint coupling. Unless fractional-microsecond pulses (order 0.1μsec) are at a very low impedance level, these pulses will normally feed loads that are predominately capacitive. When relatively large amplitudes and small delays take precedence over pulse shape, it is best to replace the load resistor of the transformer with an inductance; the inductance is normally necessary to produce overshoot on the pulse and thereby permit the use of a crystal diode to damp out oscillations. Such circuits have handled 0.1 μsec pulses at repetition frequencies as high as 3 megacycles without appreciable variation in pulse amplitude as a function of repetition frequency.

There is a good possibility that the present Whirlwind transformers designed for 0.1 μsec pulses could be improved upon; this is particularly true when the transformers drive on RLC load instead of a predominately resistive load. It is believed that non-metallic magnetic materials (ferrites) might prove to be better than metallic materials for cores of transformers designed for 0.1 μsec. pulses.

A. Introduction

Information in Whirlwind I is transmitted in the form of unilateral pulses. The polarity of a pulse is reversed every time it passes through a vacuum-tube amplifier. If capacitive coupling is used between stages, every other stage must be normally on. The use of normally-on stages is undesirable for two reasons: (1) since pulses are present less than half the time, the power consumption is greater than with normally-off stages, (2) small noise pulses and pulse overshoot are amplified. The use of inverter pulse transformers with the proper turns ratio for interstage coupling permits all pulse amplifiers to be cut off during the absence of pulses. This method of coupling is used extensively in WWI allowing the use of smaller tube types and providing noise reduction throughout the system.

B. <u>The Transformer Vs the Cathode Follower</u>

When a buffer amplifier is required to drive a low-impedance source, there may be a question of using a cathode follower instead of a conventional amplifier and a step-down transformer. The gain of a cathode follower having a load resistor R is:

$$\frac{\mu R}{r_p + (\mu+1)R} \; .$$

The gain of an amplifier using a pulse transformer with a turns ratio of N:1 is

$$\frac{N\mu R}{r_p + N^2 R} \; .$$

The gain of the amplifier exceeds the gain of the cathode follower if $N\left[r_p + (\mu+1)R\right] > r_p + N^2R$, or the gain of an amplifier exceeds the gain of a cathode follower if $N < (\mu+1)$. A cathode follower has a lower input capacitance and a lower output impedance than an amplifier so that pulses with faster rise times may be handled; however, in most cases the desired rise times and greater amplitudes can be attained with buffer amplifiers and pulse transformers.

Interstage coupling can be classed into four groups (1) point-to-point coupling, (2) multipoint-to-point coupling, (3) point-to-multipoint coupling, (4) multipoint-to-multipoint coupling. The coupling of short pulses (1 μsec or less) is accomplished in the computer by using pulse transformers of the required ratios and by the use of crystal diodes for mixing and isolation purposes.

Five transformers have been designed for use in Whirlwind circuits. The design specifications for these transformers are included in the standards books for Project Whirlwind. Pulse transformers with turns ratios of 1:1, 3:1, and 5:1 have been designed for 0.1 μsec half-sine wave pulses. A pulse transformer with a turns ratio of 5:1 has been designed for a trapezoidal pulse with a rise time of 0.1 microsecond and a duration of 1/2 microsecond. A three-winding pulse transformer has been designed for blocking oscillators. Report R-122 contains valuable information on the theory and design of low-power pulse transformers. A large portion of Volume 5 of the Radiation Lab. Series is devoted to the theory and design of pulse transformers.

C. Point-to-Point Coupling

Point-to-point coupling may be from a pulse amplifier to an adjacent stage or from a pulse amplifier to a remote stage. A one-to-one inverter is used to couple adjacent stages, and a step-down step-up arrangement is used to couple remote stages.

1. One-to-One Inverter

The 6-193-6 (Whirlwind Spec.) transformer was designed for 0.1-microsecond pulses and has a turns ratio of 1:1. This transformer is used to reverse the polarity of pulses that are coupled from a gate tube to a buffer amplifier or from one buffer amplifier to another in the manner indicated in Figure 1. The 6-193-6 transformer was originally designed to work into a resistive load of 1000 ohms; a 1000-ohm load gives the best transient response for this pulse transformer when it is driven from a current source. Because of the stray capacitances in the circuit it is impossible to present a resistive load to the pulse transformer without sacrificing a large portion of the pulse amplitude. The circuit in Figure 1 is the most satisfactory arrangement for coupling between two pulse amplifiers, assuming the pulse amplifiers are normally cut off. The transformer works into a parallel R-L-C combination that behaves something like an R-L-C peaker that is underdamped for the first half-cycle of oscillation and critically damped for the second half-cycle. The leakage inductance of the transformer and the capacitances of the transformer and the circuit combine to produce ringing on the trailing edge of the pulse. If the value of the tail-reversing inductance L is too large the ringing will not swing below the voltage axis and it can not be damped out by the resistance R. If the value of L is too small, the duration and amplitude of the pulse are decreased. A suitable value for L can be determined rapidly by experimental methods. The value of the inductance L depends on the transformer, the pulse duration and the stray capacitances of the circuit; a small value of L is required for large shunt capacitance. The shunt capacitance C_s and C_p usually total between 30 to 40 μμf in Whirlwind circuits; if the total shunt capacitance exclusive of the transformer capacitance is between 30 to 40 μμf and the pulse duration is near 0.1-microseconds a 50-microhenry inductance is satisfactory for use with the 6-193-6 transformer. The value of R is important if the circuit is to handle pulses at a high repetition rate. The value of R should be chosen to give critical damping. When R is too small the circuit is overdamped, and when R is too large the circuit is underdamped. For most Whirlwind circuits in which a one-to-one transformer is used, a resistance of 470-ohms is suitable for R. (The forward resistance of the germanium crystal diodes is about 100-ohms.) The circuit of Figure 1 will work for 0.1-microseconds pulses at a repetition rate of well over 2 megacycles without appreciable prf sensitivity.

Since most of the Whirlwind circuits that require a one-to-one inverter have nearly the same shunt capacitance, it seems that it ought to be possible to design a transformer with a lower magnetizing inductance so that a tail-reversing inductance is unnecessary. If the number of turns on the transformer are reduced, the magnetizing inductance, leakage inductance, and transformer capacitance are reduced. Pulse transformers with fewer turns were tried in an attempt to eliminate the necessity for the tail-reversing inductance. However, the results were unsatisfactory; it seems that if the number of turns on the transformer was reduced so that no tail-reversing inductance was required the pulse amplitude was less than with the 6-193-6 transformer and a tail-reversing inductance. The reason for this is not understood. One explanation could be the non-linearity of the magnetizing inductance of a transformer wound on a laminated cone. The effective permeability of an iron core varies with induction and time. The permeability of a core increases as the time for short pulses because the initial eddy currents in the core are large and tend to prevent the magnetizing flux from being uniformly distributed in the core.

2. Step-Down Step-Up

In some cases one pulse amplifier is required to drive another amplifier at a remote point. A low-impedance line is usually necessary in connecting remote points to prevent excessive distortion, attenuation, and delay of the pulses. A step-down transformer may be used to drive the low-impedance line; if the amplitude of the pulse on the line is not large enough to drive the second stage, a stepup transformer may be useful at the receiving end. This method of interstage coupling is similar to using a 1:1 transformer between stages that arennot remotely connected; a tail-reversing inductance and a damping diode are usually required. The 6-193-7 transformer is designed for 0.1-microsecond pulses and has a turns ratio of 3:1, this transformer is useful for step-down step-up purposes if the impedance of the transmission line is around 100-ohms. Again, when pulses with fast rise times (0.05-microsecond) are used it is practically impossible to terminate the 90-ohm line or the transformers with resistive loads without sacrificing considerable amplitude because of the input capacitance of the succeeding stage. As in the case of the 1:1 inverter it is usually best to leave the transformer as well as the transmission line unterminated. The step-down step-up arrangement shown in Figure 2 is useful for coupling 0.1-microsecond pulses between two stages (7AK7, 7AD7, 6Y6 etc.) connected by a length of RG 62U transmission line. (The characteristic impedance of RG 62U line is 93 ohms. The delay of RG 62U is approximately 1-microsecond per 1000 feet.) The sizes of the tail-reversing inductance and the damping resistor are about the same as for the 1:1 transformer; they can be determined best and most rapidly by experimental methods. Since the line is unterminated, reflections do occur; however, if the delay of the transmission line is less than a quarter of the delay of the pulse these reflections are damped out (because of losses in the transformer) before the overshoot of the pulse is completed and do not cause any difficulty. If the input capacitance to the succeeding stage is small a transformer with a high turns ratio might be useful at the receiving end; however, in most cases in Whirlwind, an increase in the turns ratio of the second transformer increases the rise and fall time of the circuit so that there is little or no gain in the pulse amplitude.

3. Effect of Transformers on Pulse Shape

Many of the circuits in the Whirlwind computer are designed for 0.1-microsecond half-sine-wave pulses. If the pulse duration is too long, the circuits do not have sufficient time to recover for high speed operation; if the pulse duration is too short, unnecessarily large tubes are required to deliver the required pulse amplitudes. The shape, amplitude, and delay of pulses is affected by the pulse transformer and their associated circuits. Usually a pulse amplifier is biased a few volts beyond cutoff to prevent small noise signals from being amplified; since the amplifier is biased beyond cutoff the effective duration of the input pulse is decreased. If the pulse duration is to remain unchanged the plate circuitry must act to broaden the pulse: the leakage inductance of the transformer, the shunt capacitances in the plate circuit, and the tail reversing inductance affect the shape and amplitudes of the pulses out of a pulse amplifier. Because most of the stages are biased beyond cutoff and because the circuitry affects the shape of the pulses, output pulse shape becomes reasonably independent of the input pulse shape after passing through four or five similar pulse amplifiers. The analysis in Engineering Note E-138 indicates some of the effects of pulse transformers on the shape and amplitude of the pulses. In order to make any reasonably simple analysis of these pulses circuits a number of assumptions have to be made; the analysis in E-138 does not account for non-linearities in the control grid to plate transfer characteristics, the magnetizing inductance of the transformer, or the input impedance to the second tube. The analysis was primarily for a 1:1 inverter used in the manner indicated in Figure 1; however, the behavior for step-down step-up arrangements is nearly the same. The equivalent circuit of the plate circuit was simplified as shown in Figure 3.

L_1 - leakage inductance of transformer

C_1 - output capacitance of first stage

C_2 - input capacitance of second stage plus shunt capacitance of the pulse transformer

L_2 - tail reversing inductance in parallel with magnetizing inductance of pulse transformer

If L_2 is neglected and an impulse of current is supplied across C_1 the expression for e_2 is:

$$e_2 = \frac{1}{C_1 + C_2} (1 - \cos \alpha t)$$

where

$$= \frac{C_1 + C_2}{L_1 C_1 C_2}$$

It is argued that the effective duration of the input pulse is always appreicably less that the duration of the output pulse so that the idea of an impulse is usable. The expression indicates that L_1 does not influence the amplitude of the pulse but that it does influence the frequency of oscillation in the output. It also indicates that the oscillations will not swing below the voltage axis if L_2 is infinite. If the effect of L_2 is considered in the circuit of Figure 3 for an impulse of current, the expression for e_2 becomes more complex and is rather meaningless in algebraic form. However, the addition of L_1 results in the two superimposed oscillations of different frequencies so that

$$e_2 = A(\cos \omega t - \cos \beta t),$$

where A, ω and β are constants determined by the circuit parameters. From E-138 and experimental evidence it is known that the pulse duration can be lengthened by increasing either the leakage inductance of the transformer, or the shunt capacitance across the transformer, or both; in any case, an increase in pulse duration by the circuitry results in a more triangular pulse shape and/or a decrease in pulse amplitude and an increase in delay time through the circuit. The analysis in E-138 indicates that the leakage inductance of the pulse transformer affects the pulse amplitude much less than the shunt capacitance. The leakage inductance of the transformer is increased and the interwinding capacitance decreased if the spacing between the windings is increased; when the shunt capacitance is decreased the amplitude of the ringing increases and a larger tail reversing inductance may be used. Thus the pulse duration can be increased somewhat without a serious loss in amplitude; however, the pulse shape becomes more triangular and the delay is increased. The 6-193-6 and 6-193-7 transformers in Whirlwind circuits usually result in a decrease in pulse width from the standard 0.1-microsecond to a 0.07 to 0.08-microsecond. It is questionable whether the transformer should be designed to provide greater pulse widths; the amplitude of the pulse at its final destination is usually the deciding factor that must be considered, assuming the time delay is not too great.

The point-to-point coupling in Whirlwind circuits is primarily a matter of attaining moderate Q's and high L-C ratios. Normally higher Q's for transformer can be attained by using cores of magnetic materials; however, as the pulse duration is decreased the effective premeability of a core must be sacrificed in order to keep the eddy circuit losses at the same order of magnitude as the copper losses (ideally core losses should equal copper losses for highest Q's). In general, pulse transformers have been designed to drive resistive loads; in these cases, eddy-current losses were not so important as long as they were small in comparison to the power dissipated in the load resistor. The leakage inductance and shunt capacitance of a transformer increase with the number of turns; in order to attain a minimum transformer capacitance and leakage inductance a high permeability core is desired to reduce the number of turns required for a desired magnetizing inductance. Experiments with the aircore pulse transformers in Whirlwind circuits has shown that these transformers are comparable in performance to the 6-193-6 and 6-193-7 transformers. Although a greater number of turns are required for the air-core transformers so that the leakage inductance and interwinding capacitance are increased, the performance of air-cored transformers is nearly as good as the ones with Hipersil cores because (1) the leakage inductance does not seriously affect the pulse amplitudes unless the grid of the succeeding stages is driven positive (2) the shunt capacitance of air-cored transformer is not necessarily as great as the ones with Hipersil cores (3) the core losses with Hipersil cores are not negligible. The effective shunt capacitance of transformers with Hipersil cores is affected by the winding-to-core capacitance and the core-to-ground capacitance as well as the interwinding capacitance. If air-cored transformers work nearly as well as those with Hipersil cores it may be possible to design transformers with powdered iron cores that work just as well, if not better. Non-metallic magnetic materials (Ferroxcubes, Ferramics) have been developed that have effective resistivities as good as powdered iron cores and greater effective permeabilities than powdered iron cores; these materials are being produced in small quantities by North American Phillips Co. and General Ceramics & Steatite Co. The new magnetic materials may make possible better transformer designs for short pulses and relatively large capacitive loads.

C. Multipoint-to-Point Coupling

Multipoint-to-point coupling is similar to point-to-point coupling; however, crystal diodes are useful to mix the signals from the various sources at the receiving point. The circuit of Figure 4 shows how several signals may be mixed at a common point. The mixing diodes prevent a signal on one line from feeding back on to adjacent lines, and also, these diodes prevent the amplitude of the pulse on one line from being affected by the loading effort of the adjacent lines and the magnetizing inductances of the adjacent transformers.

D. Point-to-Multipoint Coupling

Point-to-multipoint coupling is accomplished by driving a low-impedance line with a buffer amplifier and tapping the line at the desired receiving points as is indicated in Figure 5. The input impedance of the receiving points should be high compared to the characteristic impedance of the line so that the pulse amplitude at each receiving point is nearly the same. Generally speaking, step-up transformers cannot be used at the receiving points because the increased load on the transmission line results in excessive attenuation.

Since the clamping circuits at the receiving points will clamp to any pulse overshoots, it is important that the overshoot of the pulse should be small. The magnetizing inductance should be large to reduce the overshoot; however, the leakage inductance and capacitances should be small to prevent excessive attenuation and distortion of the pulse. The turns ratio of the pulse transformer should be chosen so that the amplitude of the pulse on the line is a maximum; if the pulse amplifier has linear characteristics, the voltage gain is a maximum if $N^2 Z_o = r_p$.

where N = turns ratio of transformer

Z_o = line impedance

r_p = plate resistance of tube

An increase in turns ratio increases the resistive load that appears in the plate circuit while decreasing the effect of capacitance on the line side of the transformer. If the turns ratio is too large the R-C time constant is increased to the point where the rise time of the circuit is excessive and possibly there is a loss in pulse amplitude; if the turns ratio is too small the low impedance in the plate circuit results in a loss of pulse amplitude. A turns ratio of 3:1 is about optimum for a transformer driving a 90-ohm load with 0.1-microsecond pulses from tubes having an output capacitance of 10-15 micromicrofarad; the 6-193-7 transformer was designed for 0.1-microsecond pulses and has a turns ratio of 3:1. Since the pulse amplifier driving the lines is usually biased below cutoff, the plate circuit must broaden the pulse if the pulse duration is to remain close to the standard width; the best method of increasing the pulse width without seriously affecting the pulse amplitude is to increase the turns ratio of the transformer.

For some cases in point-to-multipoint coupling an amplifier is required to drive a low impedance line in two directions; this is desirable on long lines feeding several points, in order to prevent excessive attenuation on the line. If a 90-ohm line is driven at its midpoint a 5:1 transformer is more desirable than a 3:1 transformer. The 6-193-8 transformer has a turns ratio of 5:1 and is designed for 0.1 microsecond pulses. The higher turns ratio helps to compensate for the decrease in pulse width resulting from the amplifier being biased below cut off. An increase in the rise time of the output pulse makes termination problems on the line simpler.

The 6-193-10 transformer is designed for trapedzoidal pulses with a rise time of 0.1-microsecond and a duration of 0.5-microsecond; this transformer has a 5:1 turns ratio and is designed to drive impedance of from 50 to 90 ohms. The magnetizing inductance of this transformer must be large enough to preserve the flat portion of the pulse and to prevent excessive overshoot; the leakage inductance and shunt capacitance of the transformer must be small enough to preserve the 0.1-microsecond rise time of the pulse.

If the overshoot of pulses passing through a transformer is small the recovery time is necessarily large because a transformer cannot pass a d-c component; this means that the voltage baseline will gradually shift an amount equal to the d-c component of a chain of pulses at a high repetition rate so that the effective amplitude of the pulse is reduced. Since computer circuits must handle pulses over a wide range of repetition frequencies, any form of prf sensitivity is undesirable. There are two possible methods of reducing prf sensitivity resulting from the averaging effect of the transformer: (1) the magnetizing inductance may be increased so that there is no appreciable shift in the base line for the longest chain of pulses at the highest repetition frequency to be used. This method of reducing prf sensitivity will work well if the chains of pulses are not too long, and if there is sufficient recovery time between the chains of pulses. (2) A diode may be used in the manner indicated in Figure 6. The action of the diode in this circuit is analagous to the action of the diode in a clamping circuit for capacitively coupling unilateral pulses; the capacitor in a clamping circuit discharges slowly during a pulse and recharges rapidly after the pulse has ended.

In the circuit of Figure 6 current builds up comparitively slow in the transformer winding while plate current flows but decays rapidly after plate current ceases and the pulse overshoots. In other words the time constant is L/R: R is equal to the load plus the forward resistance of the diode during the pulse and to the back resistance plus the load during the pulse overshoot. The overshoot of the transformer is large with the diode but since the back resistance of the diode is large compared to the line impedance this overshoot does not appear on the line. The 6-193-7 transformer used in the manner indicated by Figure 6 can handle 0.1-microsecond pulses at a repetition rate as high as 3-megacycles without appreciable prf sensitivity. The principal disadvantage of the circuit of Figure 6 is the loss in amplitude across the forward resistance of the diode. The D359 diode has an exceptionally low forward resistance (about 30-ohm)and a high current carrying capacity (500 ma peak). If a transformer is used to drive a 90-ohm line and the forward resistance of the diode is 30-ohms there is a 25% loss in amplitude. If pulse amplitude is of prime importance, the voltage drop across the diode may not be tolerable; however, if a chain of high prf pulses are desired with very little prf sensitivity (in test equipment for example) the circuit of Figure 6 is very useful.

E. Multipoint-to-Multipoint Coupling

Multipoint-to-multipoint coupling is illustrated by Figure 7; multipoint-to-multipoint coupling is simlar to point-to-multipoint coupling except that the driving amplifiers have to be isolated from the line to prevent excessive attenuation. Each line driver must be capable of driving a line in two directions. The line drivers are isolated from the line when not in use by crystal diodes used in the manner indicated by Figure 5. The receiving amplifiers are coupled to the line in the manner indicated by Figure 4.

F. Conclusion

If a coupling transformer is designed to drive a load that behaves nearly like a pure resistance, the design procedures in report R-122 are directly applicable. The core material of the transformer is, of course, a very important factor to consider in design. A core material should be chosen that has the highest effective incremental permeability for the pulses the transformer is required to pass. It is costly to wind transformers on stacked cores and continuously-wound cores; for this reason Mu metal and Permalloys may not be chosen as core materials even though these materials may have a somewhat higher effective permeability than materials that are manufactured into two-pien cores. In any case, better pulse transformers can be designed for resistive loads if core materials with higher effective incremental permeabilities become available.

The design of transformers to drive capacitive loads is not readily apparent from report R-122, although the general concepts do apply. Actually, not too much thought has been given to the design of pulse transformers for capacitive loads. If the transformer driving a capacitive load is required to pass a trapezoidal pulse with fidelty, the transformer circuit will have to have a low Q, and the design of the transformer will be very similar to the design of transformers for resistive loads. Section 3.43 of R-122 discusses a wide-band transformer circuit that is useful for capacitive loads. In the point-to-point coupling for Whirlwind circuits pulse amplitude, pulse delay, and recovery time take precedence over pulse shape; these coupling circuits are essentially peaking circuits so that high Q's and high L/C ratios are desired.

With improved magnetic materials for high frequencies, it seems likely that better transformers can be designed for coupling circuits that are essentially peaking circuits.

Signed: C A Rowland
C. A. Rowland

Approved: N H Taylor
N. H. Taylor

CAR:sst

Drawings: A-35261
A-35262
A-35263
A-35264

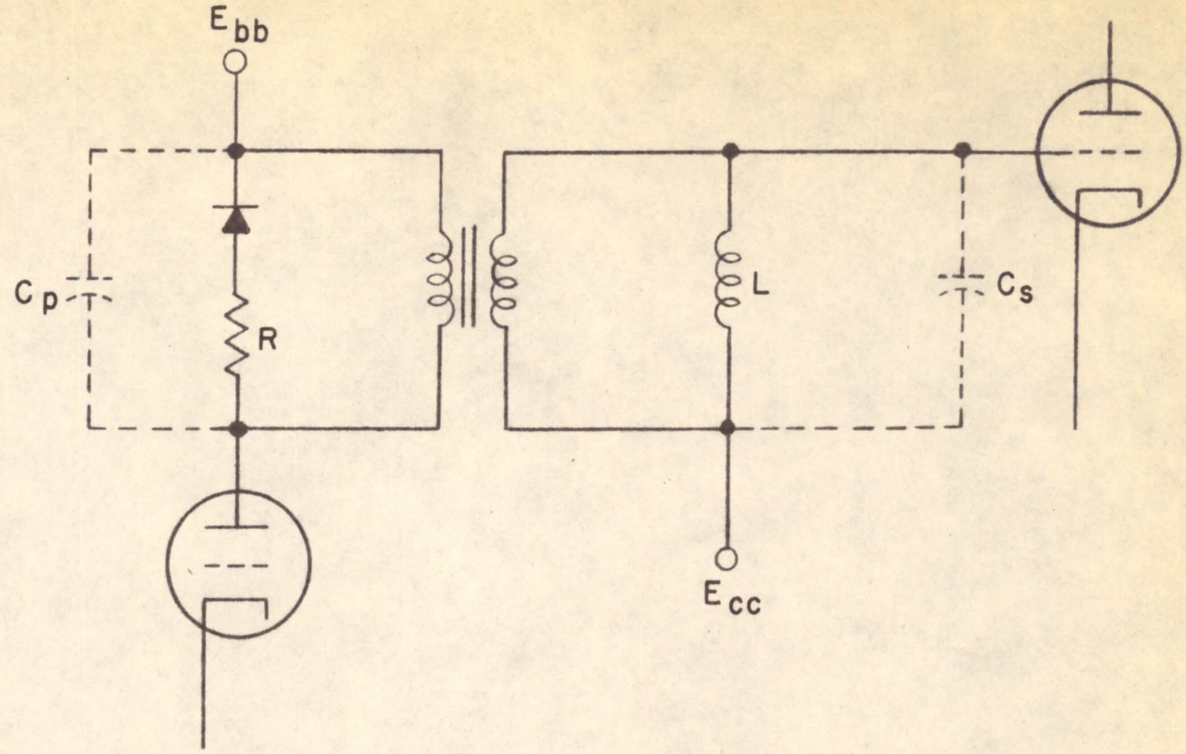


FIG. 1

COUPLING ARRANGEMENT BETWEEN ADJACENT STAGES.

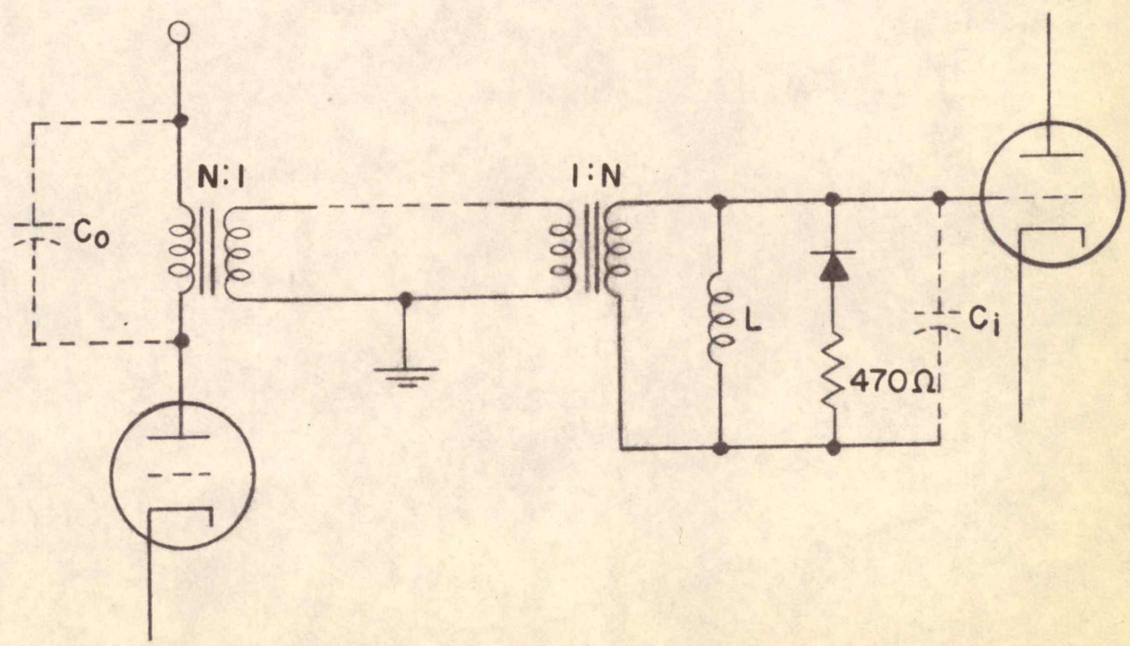


FIG. 2

COUPLING BETWEEN REMOTE STAGES

A-35261

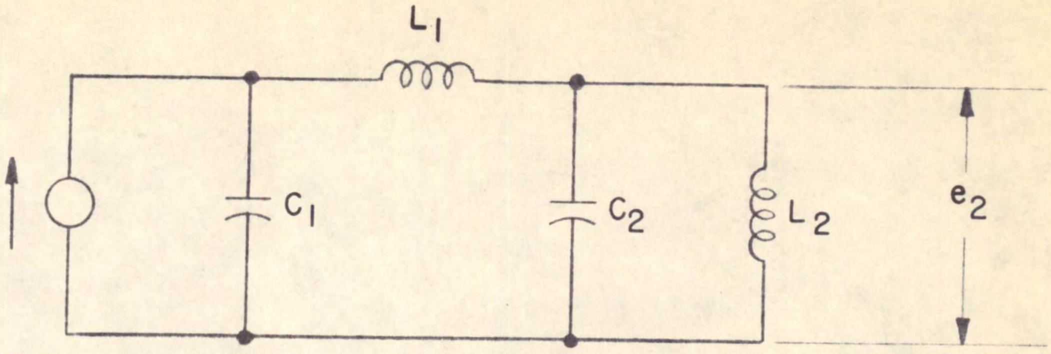


FIG. 3

SIMPLIFIED EQUIVALENT CIRCUIT FOR TRANSFORMER COUPLING

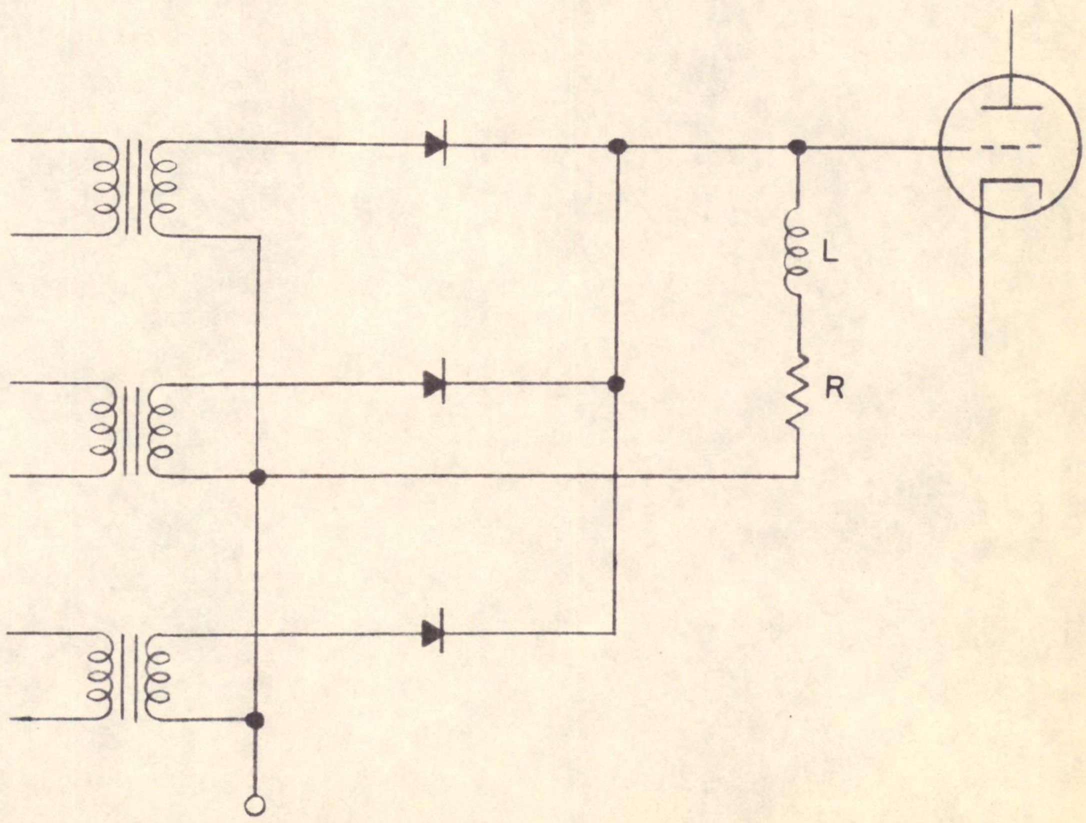


FIG. 4

MULTIPOINT TO POINT COUPLING

A - 3526

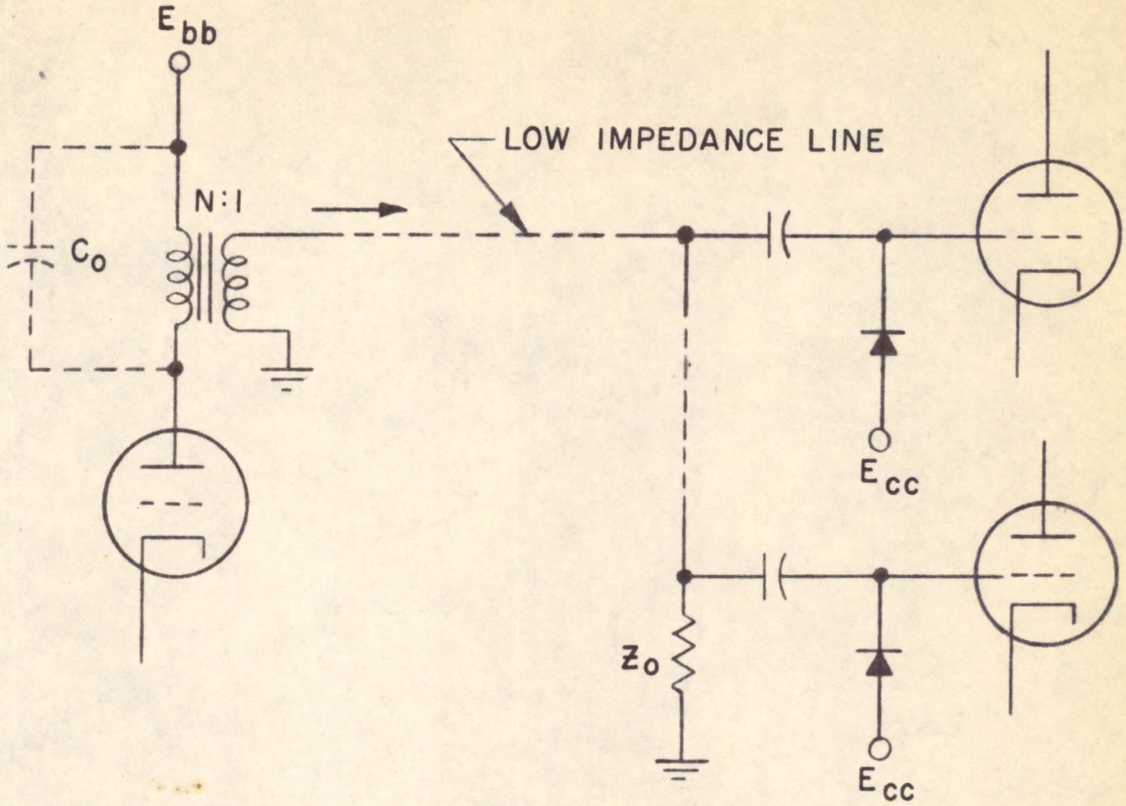


FIG. 5

POINT TO MULTIPOINT COUPLING

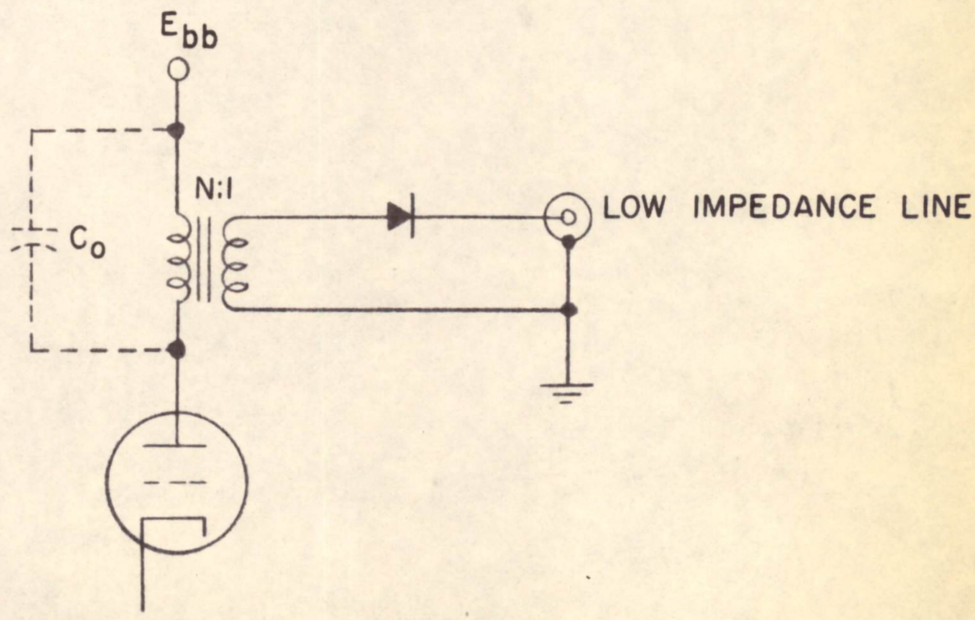


FIG. 6

THE USE OF AN ISOLATING DIODE FOR TRANSFORMER COUPLING

A-35263

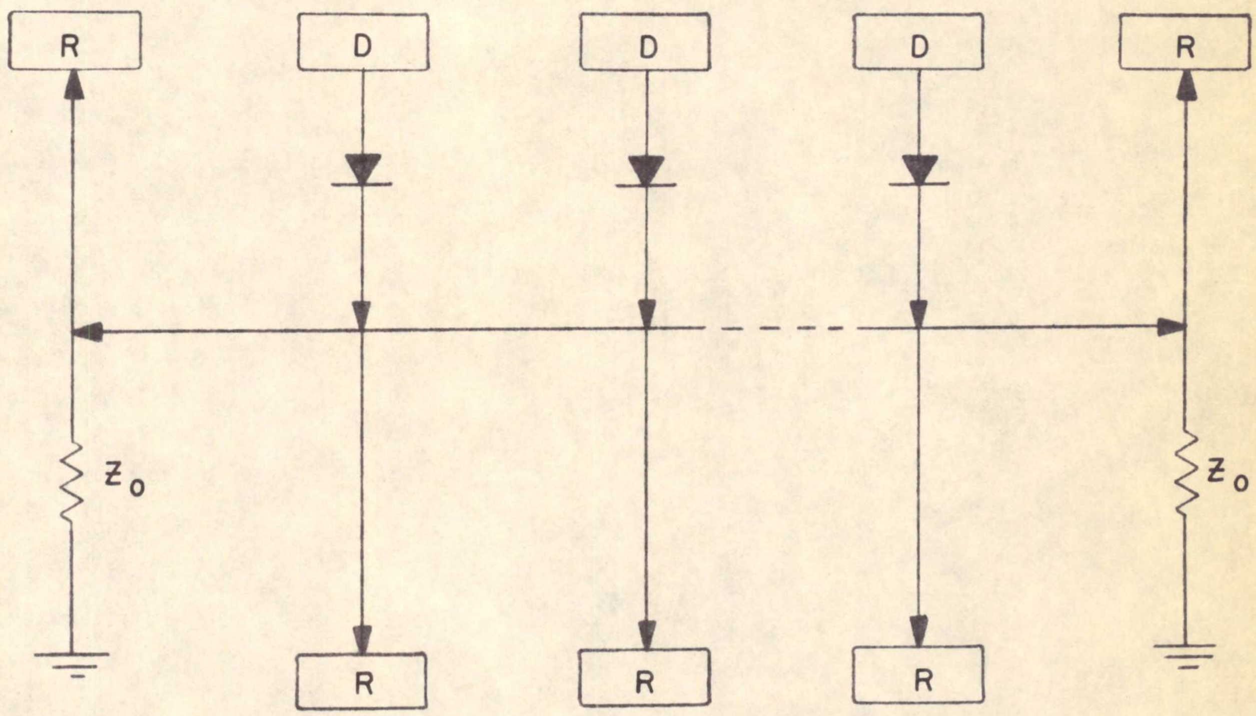


FIG. 7

MULTIPOINT TO MULTIPOINT COUPLING

Project Whirlwind
Servomechanisms Laboratory
Massachusetts Institute of Technology
Cambridge, Massachusetts

SUBJECT: TECHNIQUES FOR USING STANDARD AUTOMATIC SUBROUTINES

To: Mathematics Group

From: C. W. Adams

Date: February 10, 1950

Abstract: Standard automatic subroutines are programs for the evaluation of frequently needed functions or for the performance of routine computational chores. Such subroutines, which are intended primarily for use as a part of a larger main program, will be kept on permanent file in the slow-speed memory of a computer and will be inserted into high-speed memory along with any program of which they are to be a part. A subroutine, if it is to be referred to from several different parts of a main program, must be prepared for each use by the insertion of new addresses into some if its orders.

In the present note, various preparation techniques are discussed. Most promising is a deferred preparation, in which the only preparatory order needed in the main program is a transfer of control order, all of the necessary changes in the subroutine being performed by orders in the subroutine.

In the Whirlwind computer the address of the register next after the register containing a transfer of control is stored automatically in the A register by the transfer of control (sp) order. This register, next after the one containing the sp order, is a prearranged location characteristic of one and only one sp order, and as such this register and the ones following it can be used to store any addresses which must be available to the subroutine in making the necessary preparations. That is, the correct return address and the addresses of registers which are to contain all necessary extra addresses can all be deduced within the subroutine with no assistance from the main program.

Introduction

The simplest computer imaginable would need to have only a few basic abilities -- in fact, the ability to (1) sense and (2) complement a selected binary digit would be sufficient. Unfortunately, many repeated applications of these basic operations would be required to perform a single complete addition since the addition of each pair of digits and each carry would have to be ordered separately; so the program for any sizable problem would be unwieldy at best. In designing a computer, therefore, one builds a control and an arithmetic element of at least enough complexity to be able to perform many of the elementary arithmetic operations in response to single orders in the program. The number of elementary operations so built in may influence the speed, effective memory capacity and convenience in use of the computer, but it in no way limits the capability of the machine. For just as addition can be programmed using only a single-digit sense-and-complement order, any of the more complicated numerical operations can be programmed using the simpler built-in operations.

The Whirlwind computer has built into it the abilities to add, subtract, multiply, divide, find magnitude, and point off (find the characteristic of the $\log_2 x$ for any given x), using numbers within the prescribed range. Other operations, such as the evaluation of $\sqrt{x}$, log x, e^x, sin x, etc., are not built in. The line between included and excluded operations is drawn, as it is in all such cases, by a compromise based on such considerations as the flexibility, convenience, speed, economy of construction, simplicity, and ease of maintenance desired; these considerations hinge in turn on the intended purpose of the machine. The decision is made less difficult by the comparative ease with which the non-automatized operations can be synthesized from the elementary ones and made to seem almost automatic.

Purpose

The intention of this report is to combine and bring up to date under one cover the pertinent information on the means of accomplishing programmed automatization of the non-elementary but nonetheless common numerical operations. In the present report, a working knowledge of coding for the Whirlwind computer is presupposed. The Appendix contains a summary guide to coding and a brief but exact description of the Whirlwind order code as of January 1950.

Standard Automatic Subroutines

Programs that are commonly used as part of a larger program are called subroutines. Because the subroutines described in this report will be so arranged that they seem to be built in, these subroutines are called automatic. Only one or at most a few of the many possible variants of the subroutines for each function will be kept permanently available (in the case of sin x, for example, different routines would be provided at least for (1) x in revolutions, (2) x in radians, (3) x in radians times some scale factor.) Such subroutines will then be called standard automatic subroutines for the evaluation of the given functions. In this note, techniques for preparing to use the subroutines are discussed at length.

Standard automatic subroutines will be kept on permanent file on the slow-speed film storage of the Whirlwind computer and each will be inserted into high-speed storage along with any program in which the particular subroutine is needed. Some indication of the necessary subroutines and of the addresses of the storage registers to be occupied by each subroutine will be given as part of each main program. That is, in writing each program, the programmer will be able to select for use in his program any of the standard automatic subroutines and he will be able to designate the storage registers in which each of the selected subroutines is to be stored, subject only to the condition that his assignments are compatible with the length of each subroutine so that there are enough consecutive storage registers available for each of the subroutines.

According to the present plans for Whirlwind I, all of the standard automatic subroutines will have been written, converted to binary form, and stored in some order on one (or more if necessary) roll of film called a library film. Each subroutine will have been written under the assumption that it is to be stored beginning at register #1024. All the necessary constants, except possibly some universal constants like 1/2, will be included as a part of the subroutine, stored in registers immediately following the last order of the subroutine. The purpose of writing each subroutine starting at register 1024 is to permit easy discrimination by the computer between those orders whose address sections refer to other parts of the subroutine and those orders whose addresses are irrelevant or refer to something else (such as a number of shifts, an address of an external device, or an address of some universal constant, all of which would normally be less than 1024).

The actual insertion of a program into the computer will probably be done with the aid of three preliminary routines, all of which will presumably be stored at the beginning of the library film and will be put into the high-speed storage of the computer by means of the ri operation. The main program to be performed will have been typed out in some normal fashion on an automatic typewriter which at the same time prepares a perforated tape. (Flexowriter equipment will be used for this purpose.) Each character (i.e., each of the 50 keys and controls on the typewriter) has a six-digit binary representation on the tape and this binary-coded information can be translated in the computer to the proper binary form and stored in the proper registers by means of a suitable conversion program. The typewritten form of each new program will also have, probably at the end, some indication of the subroutines needed and the addresses assigned to each. The conversion routine, the first of the three preliminary routines, will then turn this information, properly translated, over to the second preliminary routine, a library selection routine, which will select the desired subroutines from the library film. The third subroutine (the adaptation routine) will take each subroutine and change the addresses as necessary to adapt the subroutine to its assigned place in storage. These preliminary routines will be discussed in a subsequent note.

A library of subroutines is actually being built up, starting with the common function evaluations, most of which have already been prepared in preliminary form and published (cf. E-170, C-70, C-77 for e^x and log x, sin x and cos x, and $\sqrt{x}$ respectively). A note which will contain revised and "final" forms of these and other subroutines is also forthcoming.

Classification of Subroutines

Standard automatic subroutines can be classified according to the amount of information which must be exchanged between the subroutine and the main program each time the subroutine is used. Obviously, every subroutine must be supplied with the proper return address -- the storage address of the next order in the main program to which control is to be returned at the completion of the subroutine. Aside from the return address, many subroutines such as those for the evaluation of x, log x, e^x, sin x, etc., need only to be given the value of x and need only to supply the value of the desired function. Since the quantity x and the resulting function of x will in most cases occupy only a single register each, the simplest procedure, apparently, is for the main program to put the quantity x into the Accumulator (AC) just before transferring control to the subroutine and for the subroutine to put the resulting function of x into AC just before returning control to the main program. Thus in this case no storage address, other than the return address, needs to be exchanged. Subroutines of this type, requiring the exchange of no addresses, will be said to be <u>zero-address</u> subroutines.

Some subroutines require the exchange of some number other than the quantity x and the result. For example, the quantity x or its result may be double-length -- i.e., require two registers to accommodate it because of its magnitude or its precision or both. (Frequently, however, a two-register result such as a number and a scale factor will be obtained from a zero-address subroutine since the result can easily be stored with the number in AC and the scale factor in some predetermined register, chosen once and for all.) Or, as another example, a subroutine intended to shift the contents of AC and BR left without roundoff must be supplied with the number of shifts to be performed. In such cases it is necessary for the main program to supply an address, over and above the return address, to the subroutine -- that address being either the address at which some necessary quantity will be found or at which some result is to be stored. Subroutines of this type, requiring that one address be exchanged, will be called one-address subroutines.

Similarly, other subroutines require the exchange of two, three or more orders. Thus, finding a quotient plus a remainder requires two addresses, one at which the divisor is stored and one at which the remainder is to be stored, with the dividend being put into AC and the quotient appearing in AC. The double-length arithmetic operations (addition, subtraction, multiplication, division) generally require three addresses, one for each of two operands and one for the result, since the three quantities involved are each double-length and cannot be stored in AC. In the case of double-length numbers, three pairs of addresses, six in all, are actually required, but it is assumed that the two halves of a double length number are stored in consecutive registers so that the second address of a pair can always be deduced from the first address and is not therefore a separate piece of information. As before, the number of addresses, exclusive of the return address, characterize the subroutine so that the quotient-and-remainder program is two-address while the double length operations are three-address subroutines.

Automatization of Zero-Address Subroutines

Fundamentally a subroutine is a set of orders which is used in several different parts of a main program but which is only to be put in one place in storage. The problem of automatizing the subroutine is just the problem of how to permit the subroutine to be effectively inserted into the main program by unconditional transfers of control from the main program to the subroutine and back again. By definition, the zero-address subroutines require the exchange of no address except the return address.

Suppose, for example, that the programmer has requested that a subroutine for the calculation of the square root of a number be stored starting in register 938. Suppose further that, at the completion of the main program order stored in register 619, the quantity x is in AC and that the square root of x is wanted. Then the order stored in register 620 might be sp 938 which would transfer control to the start of the square root subroutine. At the end of the square root subroutine is another sp order, which in this case should be sp 621, to return control to the proper point in the main program. This address 621, the return address, must be supplied from somewhere. It obviously cannot be simply written in once and for all, for the subroutine will probably be referred to from several different places in the main program and the return address will differ in each case.

One way to supply the return address would be to store the return address, which is known in each case, in some predetermined register. Then in the main program, before the <u>sp 938</u> order, one could clear and add that address and transfer it, using the <u>td</u> operation, to the register containing the <u>sp</u> order at the end of the subroutine. This procedure requires at least two extra orders and one extra register of constant storage for each place in which the subroutine is to be inserted into the program. In addition it presupposes anticipation by the programmer who must be sure that the <u>ca</u> and <u>td</u> sequence is inserted in the main program before the quantity x is formed in AC, it being assumed that x is to be in AC when control is transferred to the subroutine.

	Reg.	Content	
main program	.		
	606)	ca 833	anticipated preparation
	607)	td 964	
	.		
	620)	sp 938	transfer of control
	.		
constants	.		
	833)	ri 621	stored constant (operation ri is used because its coded representation is 00000)
	.		
subroutines	.		
	938)	sf 965	desired subroutine
	.		
	964)	sp --	return of control
	965)	---	
	.		

ANTICIPATED PREPARATION FOR A ZERO-ADDRESS SUBROUTINE

This anticipated preparation of the return address is possible in almost any computer. But in Whirlwind I, the addition of one simple feature to the <u>sp</u> operation and the addition of one new operation, <u>ta</u>, has removed the need for the clumsy procedure just described. The function of these two orders is as follows:

<u>sp x</u> - Transfer control unconditionally to register x. Assuming that the <u>sp x</u> order is stored in register y, store the quantity <u>y+1</u> (which is obtained from the program counter before control is transferred) in the last 11 digit positions of AR.

<u>ta x</u> - Transfer the last 11 digits from AR to the last 11 digit positions of register x, leaving the first 5 digits unchanged.

Thus in the example of the previous paragraph, the return address 621 would be stored in AR, without affecting the contents of AC, by the <u>sp 938</u> order which was stored in register 620. Then the first order in the subroutine would be a <u>ta</u> order which would transfer the return address to the last order, the <u>sp</u> order, at the end of the subroutine. The intervening orders of the subroutine would determine the square root of the number in AC and leave the result in AC. In this way the red tape involved in using the zero-address subroutines is reduced to practically nothing. The square root of the number is found by simply inserting the order <u>sp 938</u> into the program whenever the square root of the contents of AC is wanted. Using the zero-address subroutines requires no more thought or effort than using a single operation built into the computer.

	Reg.	Content	
main program		:	< no anticipation necessary
	620)	sp 938	transfer of control
		:	
constants		:	< no constant needed
subroutines		:	
	938)	ta 965	deferred preparation
	939)	sf 966	desired subroutine
		:	
	965)	sp --	return of control
	966)	---	
		:	

DEFERRED PREPARATION FOR
A ZERO-ADDRESS SUBROUTINE

Automatization of One-Address Subroutines

A one-address subroutine can be automatized in much the same way as a zero-address subroutine. The only difference is, of course, that some storage address in addition to the return address must be made available to the subroutine from the main program. Consider, for example, the subroutine for shifting left without rounding off. It is assumed that the number to be shifted is in AC and BR when control is transferred to the subroutine and that the shifted result is to be left in AC and BR at the end of the subroutine.

The subroutine itself could be

Reg.	Content	
718)	ts 751	store content of AC
719)	sl 15	store content of BR
720)	ts 752	
721)	ca 751	shift and store the original content of AC
722)	sl n	
723)	ts 751	
724)	ca 752	shift content of BR
725)	mh 753+n	
726)	ad 751	add in shifted content of AC
727)	sp --	return of control
751)	---	
752)	---	
753)	2^{-15}	
⋮		
753+n)	2^{n-15}	
⋮		
767)	2^{-1}	

EXAMPLE: A SUB-ROUTINE TO SHIFT LEFT WITHOUT ROUNDOFF

Actually, such a subroutine would be valid only for n in the range $0 \leq n \leq 14$ and consequently the standard automatic subroutine might well be slightly more complicated and more general. The subroutine just given is, however, satisfactory as an example of various preparation techniques.

The most obvious technique is anticipated preparation of the extra address, using deferred preparation of the return address. In the example given, the number of shifts, n, must be inserted in the <u>sl n</u> order in register 722 and the address 753+n must be inserted in the mh <u>753+n</u> order in register 725. Note that although two different addresses are needed, one may readily be deduced from the other and consequently the subroutine is classed as a one-address subroutine. The return address must be inserted in register 727, but this can be handled as in the zero-address technique. The result is

	Reg.	Content	
main program	.	.	
	223)	ca 690	anticipated preparation of the extra address (es)
	224)	td 722	
	225)	ad 728	
	226)	td 725	
	.	.	
	240)	sp 717	transfer of control
	.	.	
constants	.	.	
	690)	ri n	stored constant address
	.	.	
subroutines	.	.	
	717)	ta 727	deferred preparation of return address
	718)	ts 751	desired shift-left subroutine
	719)	sl 15	
	720)	ts 752	
	721)	ca 751	
	722)	sl --	
	723)	ts 751	
	724)	ca 752	
	725)	mh --	
	726)	ad 751	
	727)	sp --	return of control
	728)	ri 753	constant address needed to prepare address for the <u>mh</u> order
	.	.	

ANTICIPATED PREPARATION OF A ONE-ADDRESS SUBROUTINE

A second technique is the use of a deferred preparation of the extra address as well as of the return address. This can be accomplished by storing the extra address in some prearranged location so that the preparation orders can be put into the subroutine, thereby obviating the duplication of the preparation orders in the main program. It is not satisfactory to simply decide that the address should be stored in some one register, such as in 690 as it was in the example; for since the address will presumably be different each time the subroutine is used, the main program would still be required to transfer the address to some given location such as 690. (Although in the case at hand, in which the quantity n is needed in two different orders, some economy of orders could indeed be made by letting the main program put only the first address into the subroutine and letting orders in the subroutine form the second address from the first.)

In the deferred preparation technique to be described, the extra address is stored in a location so chosen that each reference to the subroutine from the main program designates uniquely its own prearranged location. Thus far in this discussion, the return address has been the address of the register next after the register which contains the transfer of control (the sp order). However, this register, next after the one containing the sp order, is a prearranged location characteristic of one and only one transfer of control point, and as such this register could be used to store the necessary extra address. The correct return address would then be the address of the register second after the transfer of control point.

	Reg.	Content	
main program		⋮	< no anticipation necessary
	240)	sp 717	transfer of control
	241)	ri n	the necessary address, stored in a dummy order
		⋮	
constants		⋮	< no constants needed, except those directly associated with the subroutine
subroutines		⋮	
	717)	ta 723	preparation for deferred preparation of extra addresses
	718)	ta 733	partial deferred preparation of the return address
	719)	ts 751	storing the contents of AC and BR (start of the shift-left subroutine)
	720)	sl 15	
	721)	ts 752	
	722)	ao 733	completion of deferred preparation of the return address
	723)	ca --	deferred preparation of the extra address (es)
	724)	td 728	
	725)	ad 734	
	726)	td 731	

(continued on next page)

	Reg.	Content	
subroutines (continued)	727)	ca 751	completion of the desired shift-left subroutine
	728)	sl --	
	729)	ts 751	
	730)	ca 752	
	731)	mh --	
	732)	ad 751	
	733)	sp --	return of control
	734)	ri 753	constant address needed to prepare address for the mh order
	.		
	.		
	.		

DEFERRED PREPARATION OF A ONE-ADDRESS SUBROUTINE

Automatization of Several-Address Subroutines

The generalization to several addresses of the preparatory techniques just described for one-address subroutines is almost trivial and little need be said. The anticipated preparation works in exactly the same fashion except that more than one address must be stored in constant storage and must then be transferred by orders in the main program to the subroutine. The deferred preparation is likewise almost unchanged; the several addresses are simply stored as dummy orders in consecutive registers immediately following the sp order in the main program, and the return address is simply the address of the register next after all of the several dummy orders.

History of the Discontinued "Automatic Subprogram" Operations

One of the most important set of three-address subroutines for a computer with comparatively short register length is the double-length operation subroutines. Some time ago (cf. M-111, pps. 8-10, dated October 6, 1947) it was proposed that a set of five special operations be incorporated into the design of the Whirlwind computer primarily to facilitate work with double-length numbers. The first of these operations, designated by as, would (1) transfer the contents of AC bodily into BR and (2) transfer the as order itself into AC. Three of the operations were logically identical; these operations, designated by ax, ay, az, would (1) transfer the ax (or ay or az) order itself into AR, (2) transfer the return address from the program counter to register 2047, and (3) transfer control unconditionally to some preselected -- i.e., wired in -- storage address, three different addresses being selected, one by ax, one by ay, one by az. A fifth operation, logically almost equivalent to the present ta operation but designated by ro, permitted the contents of AR to be read into AC.

The intended function of these operations is best illustrated by an example. Suppose a double-length addition subroutine is to be used and that the augend is stored in registers 618 and 619, the addend in registers 712 and 713, and the sum is to be stored in registers 832 and 833. Then the preparatory orders would be

as 618
as 712
ax 832

The first order would put the address 618 into AC; the second order would shift the address 618 into BR and put the address 712 into AC; the third order would put the address 832 into AR, would transfer control to a preselected position (x), and would store the return address in register 2047. Then the subroutine for double-length addition, stored beginning in register x, would proceed to unstack the various addresses stored in AC, AR and BR and supply them where needed in the subroutine, would deduce from the given addresses the second address of each pair (e.g., 619 = 618 + 1), would transfer the return address from register 2047 to the return of control (sp order) at the end of the subroutine, and would then perform the double-length addition.

The so-called "automatic subprogram" operations were eliminated from the Whirlwind I order code (cf. E-235, May 6, 1949). The reason for mentioning them here is threefold. First, these operations were referred to in a number of notes on programming techniques written in 1948 and it seems advisable to take cognizance of them for the benefit of anyone who has already or may yet encounter references to them in the literature of Project Whirlwind. Second, these operations point out at least one way in which special operations can be built into the machine to facilitate particular applications. Third, the method used is fundamentally a good one and provides a standard for comparison with other techniques. The reason that the automatic subprogram operations were dropped was simply that their value did not justify their existence when compared to the possible value of other special built-in operations. It should be noticed that there would be almost no gain, in fact less than none in some cases, in using the automatic subprogram operations in preparing for zero- or one-address subroutines because (1) only three locations (x, y and z) and consequently only three different subroutines can be used and (2) the use of the arithmetic element in storing addresses is obviated by the fact that in many cases the numbers themselves can be stored in AC even more effectively than their addresses.

Evaluation of the Preparation Techniques

There are several criteria by which a programming technique should be evaluated. The most important of these criteria can be summed up in the form of four questions.

(1) Is the technique easy to learn and to use?
(2) Does the technique reduce coding and manual preparation time?
(3) Does the technique reduce the storage capacity needed to solve the problem?
(4) Does the technique reduce the computing time needed to solve the problem?

In evaluating the three methods discussed in this report, one might prepare a table of comments on their relative merits. Of course, the use of special orders is not a technique of any practical importance in the Whirlwind or any other computer at present, since the orders do not exist; but it is well to keep the possibility in mind. It should also be noted that deferred preparation of a return address is only possible in a computer such as Whirlwind having the appropriate orders (sp and ta) in its code. But deferred preparation of extra addresses is possible in any digital computer, even if the order code requires use of anticipated preparation of the return address, for once the return address is available to the subroutine, the addresses of registers containing the extra addresses are also available.

EVALUATION OF PREPARATION TECHNIQUES

criterion	anticipated preparation	deferred preparation	use of special "automatic subprogram" orders
ease in learning and applying	quite easy to learn since no special technique of coding is needed; cumbersome in use	requires special knowledge, but once learned is easy to use	requires some special knowledge, but once learned is probably the easiest to use. Lacks generality since it cannot be applied to many subroutines.
reduction in coding and manual preparation time / reduction in required storage capacity	wasteful and cumbersome	quite efficient since only the essential addresses need be inserted (including the address of the desired subroutine)	most efficient for double-length numbers, since the address of commonly used subroutines does not need to be specified by the main program, but has no advantage over deferred preparation in many applications
reduction in the required computing time (actually, use of subroutines necessarily increases computing time compared to not using subroutines at all)	uses two orders to prepare the return address, which is poor; uses two orders to prepare each distinct extra address, which is good.	uses one order to prepare the return address, which is good; uses four orders to prepare each distinct extra address, which is poor.	uses two orders to prepare the return address, which is poor; uses two, plus (to get an address out of BR), orders to prepare each distinct extra address, which is fair.

Conclusions

The deferred preparation technique will be used in connection with all standard automatic subroutines for Whirlwind I. By this method, only one order is needed to bring about the evaluation of a common function, and only one extra register is needed for each extra address which is to be supplied.

Use of the various preparation techniques is shown in the following example, where the program is not an example of efficient coding but merely illustrates the thoughtless, brute force way in which results can be obtained.

Required to evaluate $(e^x \sin y - e^y \sin x)2^8$

where it is known that $0 > x > -1$
$0 > y > -1$
$2^{-8} > e^x \sin y - e^y \sin x > -2^8$

Suppose the following subroutines are available:

Evaluation of e^x for $-1 < x \leq 0$, first order stored in register A

Evaluation of sin x for $-1 < x < 1$, first order stored in register B

Double-length subtraction, first order stored in register C

A subroutine to take a double-length number, shift it left n times and put it back in the same pair of registers, first order stored in register D

Other registers are assigned as follows:

register X contains x

register Y contains y

registers T1, T2, etc. are consecutive registers available for temporary storage

The program then is: (asterisks indicate use of a standard subroutine)

```
  ca X  }
* sp A  }  find and store e^x
  ts T1 }
  ca Y  }
* sp B  }  find sin y
```

```
  mh T1  ┐
  ts T1  │
  sl 15  ├ form and store 30-digit product e^x sin y
  ts T2  ┘
  ca Y   ┐
* sp A   ├ find and store e^y
  ts T3  ┘
  ca X   ┐
* sp B   ┘ find sin x
  mh T3  ┐
  ts T3  │
  sl 15  ├ form and store 30-digit product e^y sin x
  ts T4  ┘
* sp C     subtract the second product from the first
  ri T1  ┌ address of minuend
  ri T3  ┤ address of subtrahend
  ri T1  └ address of difference
* sp D     shift the result left 8 times
  ri 8   ┌ number of shifts required
  ri T1  └ address of operand
```

Thus it is seen that such operations as the evaluation of e^x or sin x and subtracting or shifting double-length numbers can be programmed almost as easily as if they were built in to the computer, using deferred preparation of standard automatic subroutines.

Signed C. W. Adams
C. W. Adams

Approved R. R. Everett
R. R. Everett

CWA/lfu/aec

Attached: A Short Guide to Coding.

Project Whirlwind
Servomechanisms Laboratory
Massachusetts Institute of Technology
Cambridge 39, Massachusetts

SUBJECT: CIRCUITRY, RELAY TIMING, AND OPERATION OF WWI TAPE OUTPUT EQUIPMENT

To: 6345 Engineers

From: J. S. Hanson

Date: April 10, 1951

Abstract: Self-explanatory relay timing diagrams accompanied by a brief description of relay operating cycles and circuitry problems encountered cover essential details of operation of WWI Tape Output Equipment for various output modes. A new type of timing diagram is introduced.

TABLE OF CONTENTS

I. Preliminary Cabling of Units and Switch Settings

The relay cabinet is installed in Rack TC-13 in WWI Test Control. A 36-conductor cable connects the various indicator lamp circuits to the gas tube relay register on the relay panel to accomplish data transfer from WWI to the Tape Output Equipment. A 33-conductor cable connects the relay panel with the Remote Control Unit which is located on a table in the Test Control Room along with the tape units.

The reader, printer, punch, and Remote Control Unit are cabled in accordance with attached print A-36835; cabling is relatively simple and fool-proof, since receptacles with corresponding letter and unit designations are merely connected with cables provided.

The various switches contained in the reader, printer, and punch enable the units to be used interchangeably in any of the tape preparation, tape checking, and WWI input systems without wiring changes. In order to operate these units as WWI Tape Output Equipment, switches must be set as follows.

A. Printer:

1. Keep "On-Off" switch at the left side of the printer keyboard in the "off" position until system is ready to be used.
2. Throw "Insert" switch toward the rear of the machine.
3. Always keep a sheet of paper in machine to protect platen from direct contact with type bar.

B. Reader:

1. Set "On-Off" switch (above cable receptacles) to "OFF".
2. Set "Marginal Check - Normal" switch to "NORMAL".
3. Set "Normal Stop - Read Complements" switch to "READ COMPLEMENTS".
4. Set "Normal - Clutch Control Jumped" switch to "NORMAL".
5. Place slotted metal shim under tape hold-down clamp so that sprocket teeth clear ends of slot. Be sure tape hold-down clamp is properly secured.

C. Punch:

1. Set "On-Off" switch to "ON".
2. Set right-hand switch to "NORMAL" if reproduction of feedout holes is not desired.
3. Set "Marginal Check - Normal" switch to "NORMAL".

Preferably, system power should be turned on by means of the on-off switch at the left side of the printer, otherwise the printer keyboard is deprived of the mechanical interlock which prevents the keys from being accidently depressed while the system is not in use. If this happens, a number of type bars may fly up in a tangle and jam the machine when the power is turned on at the punch.

Finally, press the "Clear" push button on the Remote Control Panel, and then press the "Start" button on the printer. The former clears the gas tube register of any random "ones" which may have resulted from previous operations, and the latter makes the tape units and printer operative, as described in Section VI, page 8. The system may then be manually stopped at any time by pressing the "stop" button on the printer. To shut down the tape output equipment, operate the power "on-off" switch on the printer.

The rotary control switch (S-1) on the relay panel must be set to correspond to the desired type of WWI data output, as shown in the tabulation "DIGIT CONNECTIONS TO SWITCH" on Dwg. R-35927-2. Table I shows the connections obtained for each of the eight switch positions:

TABLE I

Switch Position	Source of WWI Data Output to Gas Tube Register
A	Accumulator (digits 10 through 15) and Program Register (digit 9)
B	Flip-Flop Storage Register #2 (digits 10 through 15) and Program Register (digit 9)
C	Flip-Flop Storage Register #2 (digits 9 through 15)
D	Program Counter (digits 8 through 14)
E	Control Switch (digits 0 through 4) and Accumulator (digits 0 and 1)
F	Accumulator (digits 2 through 8)
G	Accumulator (digits 9 through 15)

II. "Words Only" Mode of Operation

A detailed analysis of Tape Output Equipment operation is accomplished with the aid of special timing diagrams developed by the writer and introduced here for the first time. The timing diagrams not only show the timing of every relay, contact, and electromechanical component in the system, but flow paths (cause and effect) as well. To aid in an understanding of these diagrams, a short list of symbols and their interpretations are shown at the bottom of each sheet. Entries in the left hand column include all relay coils, relay contacts, cam driven contacts, and tape unit electro-mechanical components, whether or not they are utilized in a particular mode of operation.

In the "Words Only" tape output mode, the printer records a character for each and every six-digit "word" placed in the output register by WWI. Switches on the remote control panel are set to "PRINTER ON", "PUNCH ON", "PRINT EVERY CHARACTER", and "LOCK CONTROL, WWI INPUT", and are designated on Drawing R-35927 by S8, S7, S6, and S5 respectively.

Referring now to relay timing diagram D-37308, an operating cycle is immediately initiated by a 0.1-microsecond "Print" pulse from WWI which appears at essentially the same instant that a word read into the output register causes the associated indicator-lamp circuits to drive positive the grids of the corresponding 2D21's in the gas tube register. Both of these effects are indicated on the timing diagram at 0 ms.

The "Print" pulse, stretched to approximately 4 microseconds by the blocking oscillator circuit of V-1, fires "Print" thyratron V-2 and energizes "Print" relay K-1. At 6 ms, "b" contact K-1-2 transfers the "Completion Signal" circuit from +90 volts to -30 volts. As K-1 closes in, "a" contact K-1-13 closes at 9 ms, seals in K-1 via "b" contact K-2-1, and conveniently shorts out and de-ionizes the "Print" thyratron at the same time. As "a" contact K-1-5 closes, plate voltage now applied to the relay register fires only those gas tubes with positive control grids, and energizes the corresponding register relays, (K-12 for example, which in turn closes "a" contacts K-12-5 and K-12-7 at 20 ms). Closing of remaining "a" contact K-1-6 energizes the reader clutch magnet, and after a 32-ms clutch mechanism delay, the reader cam shaft begins its rotation at 41 ms.

Closing of the "D-C Common" contact in the reader at 48 ms energizes the "translator" relays and punch selector magnets through register relay "a" contacts on the Reader Readout and Punch Readout buses respectively, and also energizes "D-C Common" relay K-2 directly. For example, with register relay K-12 closed, translator relay K-102 is energized through "a" contact K-12-5 on the Reader Readout Bus, and punch selector magnet K-202 is energized through "a" contact K-12-7. The tape punch now performs an automatic 74-ms electro-mechanical cycle of rotating the cam-shaft, actuating the selected punch mechanisms to punch the desired holes, de-energizing the clutch magnet, and advancing the tape for the next cycle.

At 55 ms, "b" contact K-2-1 opens and breaks the seal-in circuit of "a" contact K-1-13, de-energizing "Print" relay K-1. At 57 ms, "a" contact K-2-5 closes to maintain plate voltage on the relay register when "a" contact K-1-5 subsequently opens.

At 60 ms, closing of the reader "A-C Common" contact energizes the desired printer solenoid through "a" and "b" contacts on the translator relay bank, and the printer then types the corresponding character.

At 66 ms, as mentioned above, "a" contact K-1-5 opens, (relay register plate voltage is still maintained by "a" contact K-2-10), and "a" contact K-1-6 de-energizes the reader clutch magnet. Closing of "c" contact K-1-2 at 68 ms issues no "Completion" signal at this premature time because "b" contact K-2-8 is still open. Likewise, closing of the reader Feedout contact at 95 ms and its opening at 118 ms has no effect, since this contact is not used for the "Words Only" mode.

At 105 ms, opening of the "A-C Common" contact de-energizes the printer solenoid, and at 106 ms, opening of the "D-C Common" contact de-energizes the translator relays (K-102), the punch selector magnets (K-202), and "D-C Common" relay K-2.

At 137 ms, opening of "a" contact K-2-5 now removes plate voltage from the relay register (de-energizing K-12), closing of "b" contact K-2-6 sets up the "Completion" signal circuit whereupon closing of "b" contact K-12-2 in the relay register at 147 ms places +90 volts on the "Completion" signal line to terminal 12 and impresses a positive 120-volt pulse on the control grid of a 2D21 pulse generator (ref. Drawing R-33486) whose output is a 0.1-microsecond "Completion" pulse. This pulse signals WWI to clear the storage register, read into it the next word, and send a 0.1 microsecond "Print" pulse to the Tape Output Equipment to start the next cycle.

III. Relay Counter Presetting Feature for "Word-Complement" Mode of Operation

Operation of individual tape units is identical to that of the "Words Only" mode, but in addition a scale-of-two relay counter is needed to permit the printer to type a character corresponding to the particular 6-digit "Word" in the gas tube relay register during punching of the word in paper tape, and then to keep the printer inoperative during the next cycle while the complement is being punched.

In employing a relay counter, it is possible to have the counter in the wrong position at the start of a train of data because of some previous usage, an unexpected switching transient, or some manipulation of tape equipment, so that it is necessary to arrange for some method of presetting this counter immediately before recording "Word-Complement" data. Accordingly, at least one six-"zero" "Blank" signal (meaningless to the printer) must be provided by WWI as the "Blank Preset" signal. Since more than one "Blank" signal may be needed for other WWI functions, the relay counter has been arranged so that it remains in the position preset by the first "Blank" signal regardless of the number of similar "Blank" signals following.

Since operation of the tape units is essentially the same in all modes, a description of timing relations in the relay counter will be considered sufficient for the "Word-Complement" mode. For this mode, switch S6 is set to "Words and Complements".

Referring now to the "Blank" Signal Preset timing diagram (drawing D-37301), all four counter relays (K-4 through K-7) and "Blank Preset" relay K-8 are initially unenergized. As before, the preset cycle is initiated by a WWI "Print" signal which fires "Print" thyratron V-2, and energizes "Print" relay K-1 at t = 0 ms. Closing in of K-1 energizes "D-C Common" relay K-2 and reader clutch magnet K-107, and deionizes "Print" thyratron V-2. Again during the reader mechanical cycle, "D-C Common" contact S-102 closes at 42 ms and energizes "D-C Common" relay K-2, which de-energizes "Print" relay K-2 and the reader clutch magnet. Note that in this mode, there is no meaningful data yet present in the relay register because an all-zero "Blank Preset" stored in the WWI output register has kept all seven control grids of the gas tube register (V-3 through V-9) biased to -35 volts, and therefore application of plate voltage to the gas tube register at 50 ms fails to energize any of the register relays (K-11 through K-16). Hence no signal gets through translator relays and selector magnets to actuate the tape punch or printer.

Departure from a regular mode of operation now occurs at 88 ms when closing of reader "Feedout" contact S-101 energizes "Switch" relay K-4 of the relay counter through switch S6-2 which was set in the "Word-Complement" position. At 92 ms, "a" contact K-4-3 closes and energizes "Count" relay K-5 through "b" contact K-6-2, and the relay then seals in through "a" contact K-5-9.

Reopening of "Feedout" contact S-101 at 110 ms de-energizes "Switch" relay K-4, opening "a" contact K-4-3. This now has no effect on "Count" relay K-5 since it is sealed in through its own contact. However, closing of "b" contact K-4-1 energizes "Interlock Pulse" relay K-7 through "a" contact K-5-3 and the preset cycle is completed with the closing of "b" contact K-2-8 and issuance of a "Completion" signal. The equipment is now ready to punch and print the next <u>meaningful</u> character from WWI which will be a <u>word</u>, punch the following complement only, and then repeat the cycle as long as WWI continues to furnish words and complements. Relay action for this operation will be briefly discussed under Section V, page 7.

IV. Effect of Multiple "Blank Preset" Signals on Relay Counter Setting

Consider now that a second "Blank Preset" signal is furnished by WWI instead of a "Word", as shown on timing diagram D-37301. Relays K-1, K-2, and the reader perform the same functions as before, with the exception that "a" contact K-5-6 is now closed at 185 ms and 115 volts a-c is applied to "Blank or Preset" relay K-8. As the reader cycle continues, "Feedout" cam contact S-101 closes at 220 ms and energizes "Switch" relay K-4 as before, except that "b" contact K-4-1 opening at 222 ms now de-energizes "Interlock Pulse" relay K-7.

At this point it is important to note that the purpose of "Interlock Pulse" relay K-7 is to deprive "Count Interlock" relay K-6 of any voltage before "Blank or Preset" relay K-8 can drop out. This is just about accomplished, but because of the fast drop-out time of K-8, an occasional 1-ms pulse does tickle "Count Interlock" relay K-6 at 231 ms. At no time has this critical timing point given rise to operational errors either in extensive tests required for this report or during normal operation with WWI, since marginal variation of voltages only contributes to a faster drop-out of "Interlock Pulse" relay K-7 and complete disappearance of this very short pulse. The pulse would have to be of at least 5 ms duration before a drop-out of "Count" relay K-5 and the resultant tripping of the relay counter to the opposite state would be effected.

To continue with the remainder of this second "Blank" signal cycle, opening of reader "Feedout" cam contact S-101 at 242 ms de-energizes "Switch" relay K-4, closing "b" contact K-4-1 at 248 ms and restoring all relays of the counter to their original preset positions as established by the first "Blank or Preset" signal. Reclosing of "b" contact K-2-8 at 264 ms ("D-C Common" relay K-2 having been previously de-energized at 232 ms by opening of the reader "D-C Common" contact) then initiates the "Completion" signal to WWI, signifying that the Tape Output Equipment has completed a cycle and is awaiting another "Blank" signal or a "Word".

V. "Word-Complement" Mode of Operation

Referring now to timing diagram D-37303, which is a continuation of D-37301 in regards to the elapsed time scale, relay and tape unit timing in this mode is identical to that of the "Words Only" mode described in Section II, page 4, as far as 492 ms, except that now the relay counter is now part of the system by reason of the switch settings on the Remote Control unit. Tape punch and printer cycles are initiated in the normal manner at 446 and 457 ms.

At 492 ms, closing of "Feedout" cam contact S-101 now is permitted to energize "Count Interlock" relay K-6 via "a" contact K-7-9 and "b" contact K-8-7, the latter being closed because "Blank" signals needed to energize "Blank or Preset" relay K-8 are now absent.

As a result, "b" contact K-6-2 opens at 497 ms, unseals and de-energizes "Count" relay K-5. Meanwhile "Count Interlock" relay K-6 seals in at 499 ms through "a" contact K-6-10 to keep relay K-5 isolated from energizing voltage present while "a" contact K-4-3 is closed. "Interlock Pulse" relay K-7, which had dropped out at 494 ms, and "a" contact K-5-3 which opened at 507 ms (as described in Section II) now prevent "b" contact K-4-1 from de-energizing "Interlock Pulse" relay K-7 at 520 ms. All relays in the counter are de-energized, the counter is now in the "Complement" position, and the cycle is completed in the usual manner.

Upon receipt of a "Completion" signal, WWI places the "Complement" in the gas tube register and issues another "Print" signal at 543 ms. The resulting relay and tape unit operations are again the same as before, except that "a" contact K-5-6 is now open and the printer receives no 115-volt A-C signal through the Translator Relays (K-102 for example), hence no character is typed. The tape punch selector magnets are energized at 593 ms in accordance with relay register "a" contacts to punch the "Complement".

The closing of reader "Feedout" cam contact S-101 at 639 ms energizes "Switch" relay K-4, and the closing of "a" contact K-4-3 at 643 ms energizes "Count" relay K-6 via "b" contact K-6-2, and the latter seals in at 650 ms through "a" contact K-5-9. Reclosing of "b" contact K-4-1 after "Feedout" cam contact S-101 opens energizes "Interlock Pulse" relay K-7, the relay counter once again attains the "Word" position, the cycle is completed at 690 ms, and the system is ready for the next "Word".

VI. Tape Unit "Start-Stop," Punched Tape Feedout, and Delay Control Relay Functions

A. "Start-Stop"

Pressing of the "Start" button (S-109 on the reader or S-301 on the printer) is necessary in readying the tape output equipment for operation with WWI after turning on the power. Its effect on the delay control circuits are as follows: pressing of the "Start" button first opens a "b" contact to disconnect the reader clutch magnet and then closes an "a" contact which energizes Clutch Control relay K-108. A "b" contact on this relay opens the reader clutch magnet circuit at a second point, and then an "a" contact energizes the pickup coil of Delay Control relay K-109. A "c" contact of K-109 first completes the circuit to the "buck" coil of this same relay, then an "a" contact in the clutch magnet circuit closes, and finally K-109 seals in through a second "a" contact and a 4000-ohm resistor R-106.

At this point (with the "Start" button still held down) both K-108 and K-109 are closed, and there are still two breaks in the reader clutch magnet circuit: one is the "b" contact on K-108 and the other is the "b" contact on the "Start" switch. Release of the "Start" button first opens the "a" contact and de-energizes Clutch Control relay K-108 which in turn closes the "b" contact in the reader clutch magnet circuit. Finally, closing of the "Start" button "b" contact completes the circuit to the clutch magnet. Only then can "a" contact K-1-6 of the "Print" relay energize the clutch magnet, trip the clutch and start the cam contact mechanism of the tape reader, hence the necessity of pressing the "Start" button after turning on the power.

B. Punched-Tape Feedout

In the tape punch, punched tape must travel from the perforating mechanism a distance of three or four inches to the point at which it emerges from the tear-off guide, so that after completing an output tape, the "Feedout" button on the tape punch is pressed to feed out about six inches of blank tape, <u>after</u> which the tape can be torn off. Obviously, of the "Feedout" button is not used, three or four inches of data bearing tape will be left inside the tape punch when the tape is torn off.

A possibility now arises where someone might press the "Feedout" button assuming that the computer has finished reading out data, and spoil the tape. To avoid this condition, the "Feedout" button is rendered electrically inoperative by means of the "Unlock-Lock Control, WWI Input" switch when the latter is in the "LOCK CONTROL, WWI INPUT" position. The "Clear" pushbutton on the Remote Control box is likewise rendered ineffective by the same switch, since it is also possible to spoil the preparation of an output tape by inadvertently pressing the "Clear" pushbutton.

D. Delay Control Relay Function

The printer must frequently execute tabular movements of the carriage, carriage returns, shifts for capitals or symbols, and other functions which would cause it to fall behind control signals from the tape reader, so that obviously some kind of delay must be provided to make the reader pause while the printer completes a function. This is accomplished by contacts in the printer which first de-energize the reader clutch magnet and then energize "Clutch Control" relay K-108 which in turn opens a "b" contact in the clutch magnet circuit. The reader cam shaft rotates to the point at which the clutch automatically disengages and stops. The clutch magnet is not able to trip the clutch until the printer completes its function.

Energizing of "Clutch Control" relay K-108 maintains voltage on the pickup coil of "Delay Control" relay K-109 during the delay period when a-c is flowing through one of the printer machine function solenoids that actuate the keyboard. This current flow creates a voltage drop across the buck coil rectifier SR-100 and its rectified output causes "Delay Control" relay K-109 to drop out and open the reader clutch magnet circuit. By this time K-108 has had time to close and maintain the open circuit.

Completion of a machine function closes the delay contacts, de-energizes "Clutch Control" relay K-108 which then energizes Delay Control relay K-109. An "a" contact on K-109 then completes the clutch magnet circuit, whereupon the clutch is tripped and the reader starts out on the next cycle.

VII. Special Circuit Provisions

In preliminary operational tests on the breadboard version of the tape output equipment, considerable difficulty arose from excessive transients put out on both the a-c and d-c power lines with amplitudes in the order of 100 volts or more and of frequencies in the megacycle region. R-F filters provided in the printer for the centrifugally-controlled governor motor were insufficient for WWI standards so additional filtering consisting of L-8 and C-39 was provided in the 115-volt A-C power line at the point of entry into the relay cabinet.

In turning on the power to the tape output equipment, the possibility of originating a spurious "Completion" signal is eliminated by means of an R-C network consisting of R-67, R-68, and C-26 connected in the +90 and -30 volt lines immediately after the single-stage L-C filters. Condenser C-26 limits the rate of rise of voltage on the "Completion" signal line to approximately 2.4×10^4 volts per second, which is sufficiently low enough so that the pulse generator sees no more than approximately 0.8 volt at the instant the +90 volts is turned on, and can therefore produce no 0.1-microsecond "Completion" signal to interfere with other WWI operations. Note that this R-C network may be rendered ineffective if power to the Tape Output Equipment is turned on and off by means of the lever switch on the rack power control unit, since it is entirely possible that the +90 may appear before the -30, in which case a spurious "Completion" signal would result.

Large amplitude relay switching transients are eliminated from the +150 line by means of a two-stage filter containing L-4, L-5, and C-29 through C-31. Any remaining transient effects are less than WWI specifications of 1.5 volts permissible maximum.

VIII. Power Requirements, Recommended Fuse Ratings, and Circuit Conditions for Marginal Operation

The following tabulation shows the maximum a-c and d-c current inputs to the Tape Output Equipment during the various indicated modes of operation at rated nominal voltage, the fuse capacities required, and voltages resulting in marginal operation.

TABLE II

Supply Volts	Terminal No.	Max. Milliamps. One-Hole	Max. Milliamps. Six-Hole	Recommended 3AG Fuse Sizes, amperes	Minimum Voltage for Marginal Operation
-150	E2-8	0.2	0.3	1/2	(Note 2)
-30	E2-6	0.1	0.1	1/2	-1 (Note 3)
-15	E2-4	15	15	1/2	-11 (Note 4)
+90	E2-3	0.1	0.1	1/2	+40
+150	E2-7	21	140	1	+115
115 AC	--	(2.3 amperes)		3 (Note 1)	95

Note 1. Littelfuse "Slo-Blo" Cat. No. 313003, 250 volts or equivalent; all others Cat. No. 312500 (1/2 amp.) or Cat. No. 312001 (1 amp.) or equivalent.

Note 2. Margin expressed in terms of minimum voltage on terminals E1-1 through E1-8 necessary to fire relay register gas tubes. These voltages range from +170 to +187 volts.

Note 3. Equipment operates satisfactorily with almost negligible bias at this point, but disconnecting lead stops operation instantly.

Note 4. Reader shaft stops at -10 volts; all relay register gas tubes fire at -3 volts.

IX. Summary

The Tape Output System has operated dependably over a period of several months, except for usual difficulties encountered in the tape units, such as a) stalling and scorching of driving motors in tape punch and reader because of incorrect fusing and insufficient lubrication reaching the shaft bearings, b) bad arcing and burning of reader "D-C Common" cam contact, c) stalling or jamming of carriage return mechanism in printer, or d) printer governor-motor inoperative because of faulty governor contacts.

The possibility of Tape Output System failure as a result of sticking of relay armatures can be eliminated by a careful check of all plug-in Type J relays for incorrect residual and heel-piece gap settings. For the residual screw setting a gap of 1.0 ± 0.5 mil (measured between the armature and the abutting end of the relay core by means of a feeler gauge) is recommended. For the heel-piece gap (measured by feeler gage between the hinged end of the armature and the abutting end of the heel piece of relay "frame" with the armature held closed) a clearance of 1.0 ± 0.5 mil is also recommended.

Signed by *James S. Hanson*
J. S. Hanson

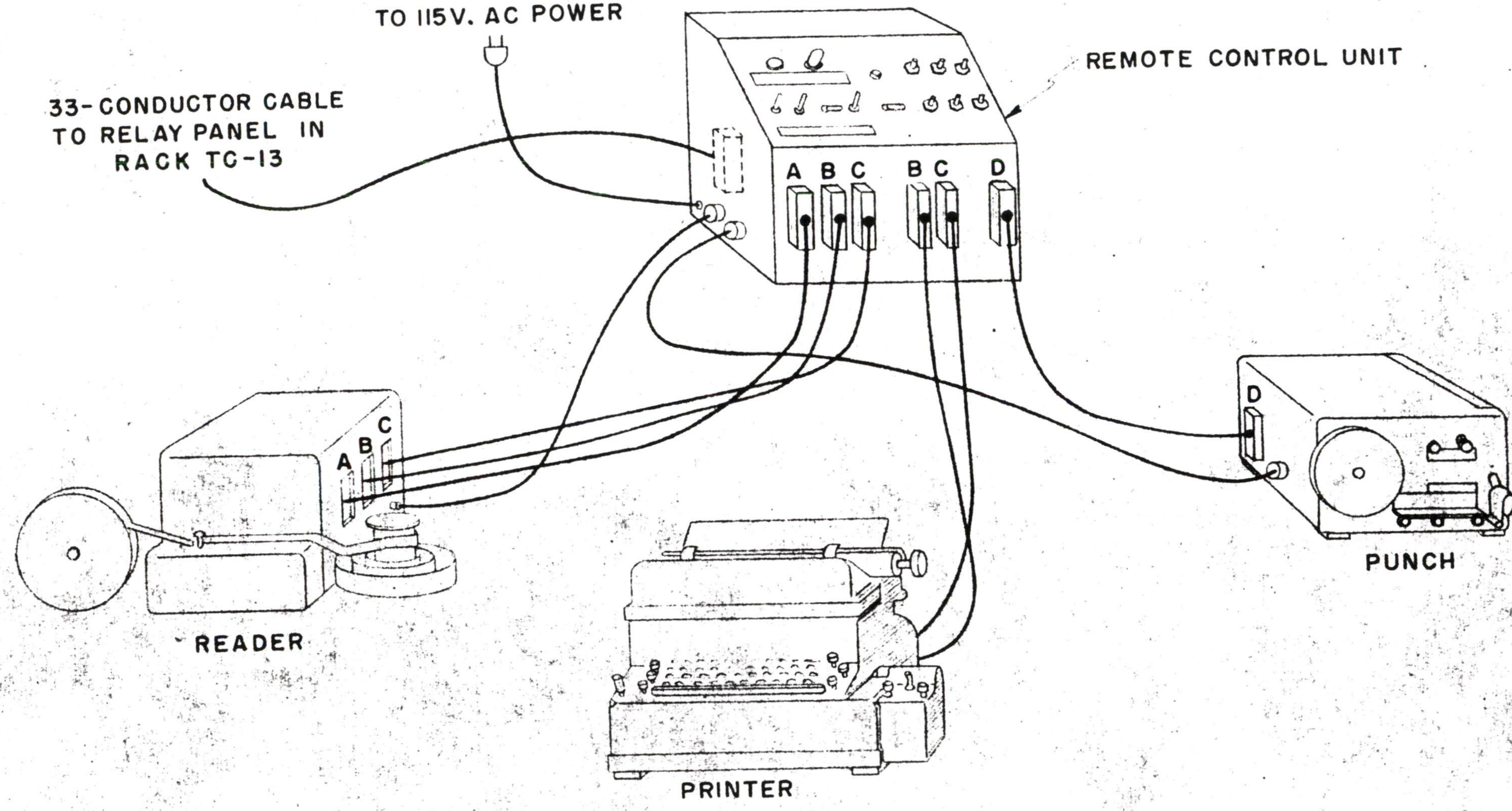
Approved by *E S Rich*
E. S. Rich

JSH:js

cc: N. Daggett
S. Dodd
R. Hunt
F. Irish
E. Rich
C. Watt

Drawings attached:

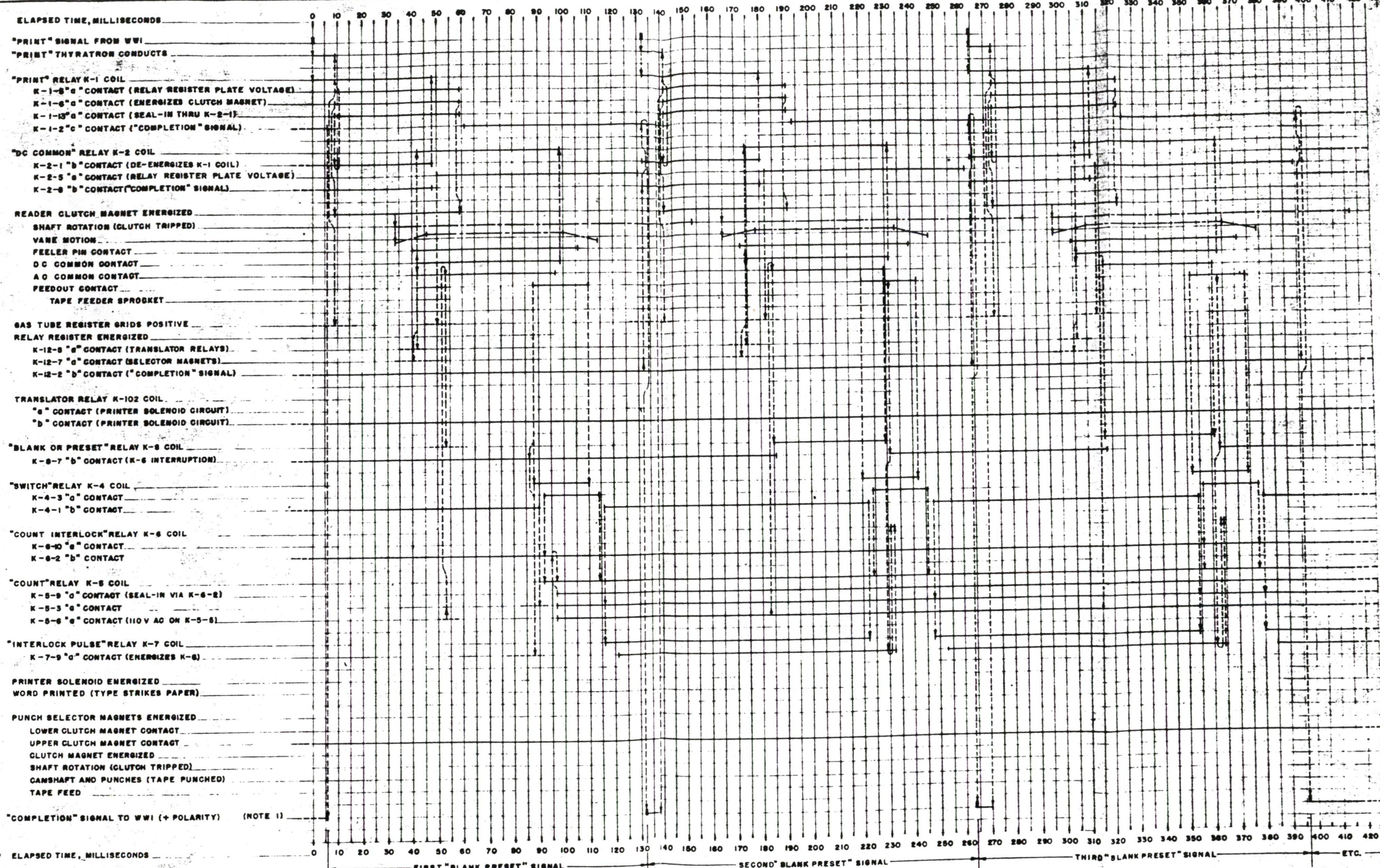
R-35927
A-36835
D-37308
D-37302
D-37303



TAPE OUTPUT SYSTEM CABLING BETWEEN READER, PRINTER, PUNCH, AND REMOTE CONTROL UNIT.

A-36835

D-37301



NOTE 1. SYMBOL (L) SIGNIFIES INSTANT AT WHICH "COMPLETION" PULSE IS RETURNED TO WWI FROM PULSE GENERATOR (REF. DWG. R-35189) HORIZONTAL LINE (———|) SIGNIFIES DURATION OF +90 VOLTAGE ON THIS CIRCUIT (TERMINAL E-1-12 ON JONES STRIP), 30V. BEING PRESENT AT ALL OTHER TIMES.

SYMBOLISM:

DENOTES PERIOD DURING WHICH A RELAY COIL IS ENERGIZED, OR AN "a" OR "b" CONTACT CLOSED.

VERTICAL DOTTED LINES INDICATE SIGNAL FLOW PATH OR AN OPERATION OF ONE CIRCUIT COMPONENT AS A RESULT OF OPERATION OF ANOTHER COMPONENT.

DENOTES BLOCKING OF SIGNAL BY OPEN CONTACT.

DENOTES SIGNAL THAT WOULD HAVE BEEN PRESENT HAD CONTACT BEEN CLOSED.

COIL / "a" CONT. / "b" CONT. / "b": COMPLETE INDICATION OF RELAY CONTACT OPERATING DELAY TIME WITH RESPECT TO RELAY COIL.

DENOTES INITIATION OF ONE MECHANICAL FUNCTION AS RESULT OF ANOTHER BY MEANS OF HORIZONTAL DOT-DASH LINES AND VERTICAL SOLID LINES, AND ALSO MECHANICAL OPERATION OF CONTACTS BY EACH FUNCTION.

K-1-13 "a" CONTACT. REFERS TO AN "a" CONTACT CONNECTED TO PIN 13 OF "PRINT" RELAY K-1, FOR EXAMPLE, AS SHOWN IN CIRCUIT SCHEMATIC R-35927.

GRADED BY / DATE
GRADE I FOR REFERENCE
GRADE II PRELIMINARY DESIGN
GRADE III FINAL DESIGN

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
DIVISION OF INDUSTRIAL COOPERATION PROJECT NO 6345

TIMING DIAGRAM FOR "BLANK" SIGNAL PRESET OF RELAY COUNTER, TAPE OUTPUT EQUIPMENT, WWI

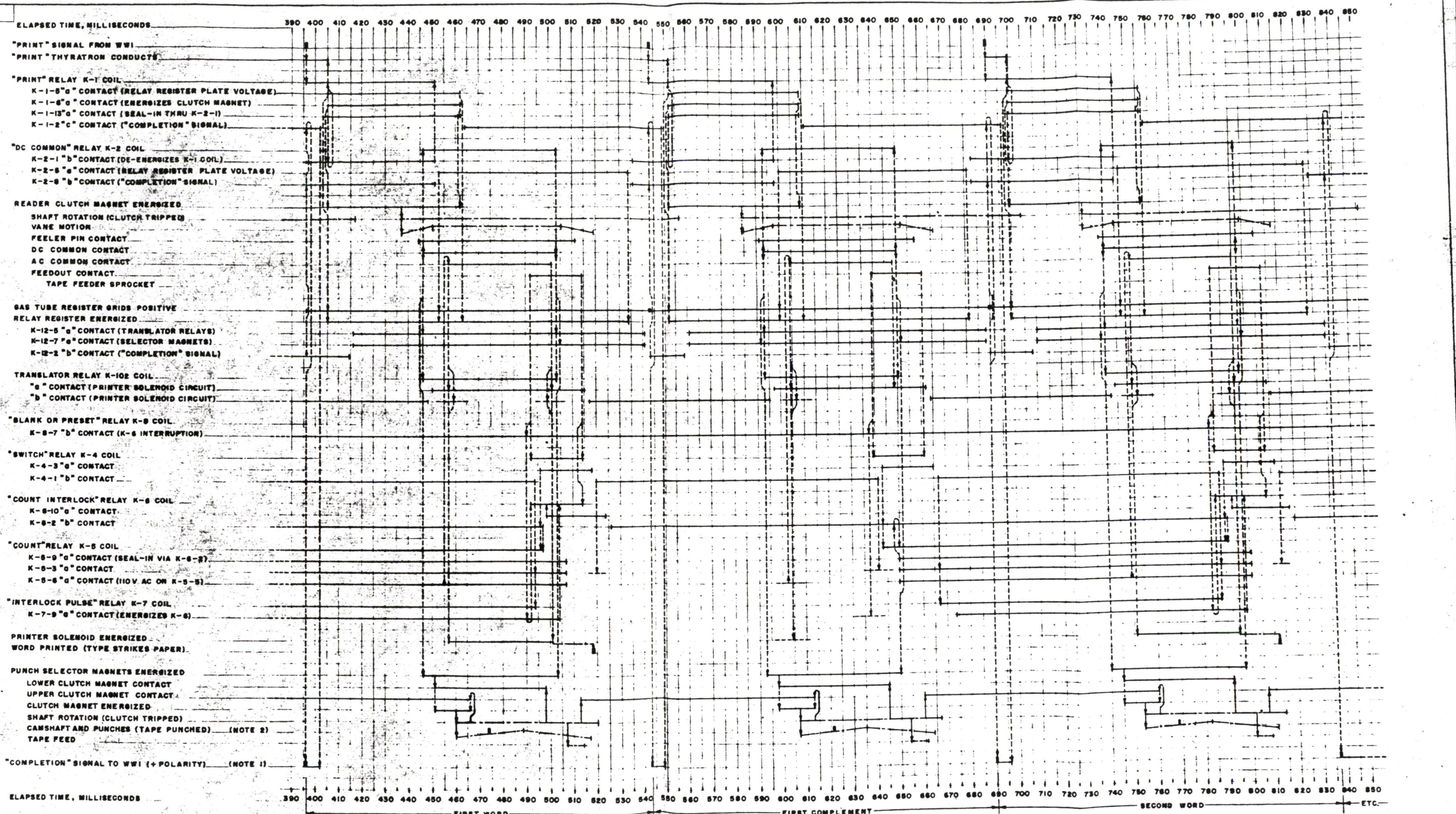
A.M.G. 3-12-51

4-9-51

D-37301

B-REDUCTION

D-37303



NOTE 1. SYMBOL (L) SIGNIFIES INSTANT AT WHICH "COMPLETION" PULSE IS RETURNED TO WWI FROM PULSE GENERATOR (REF. DWG. R-35189). HORIZONTAL LINE (———) SIGNIFIES DURATION OF +90 VOLTAGE ON THIS CIRCUIT (TERMINAL E-1-12 ON JONES STRIP), -30V. BEING PRESENT AT ALL OTHER TIMES.

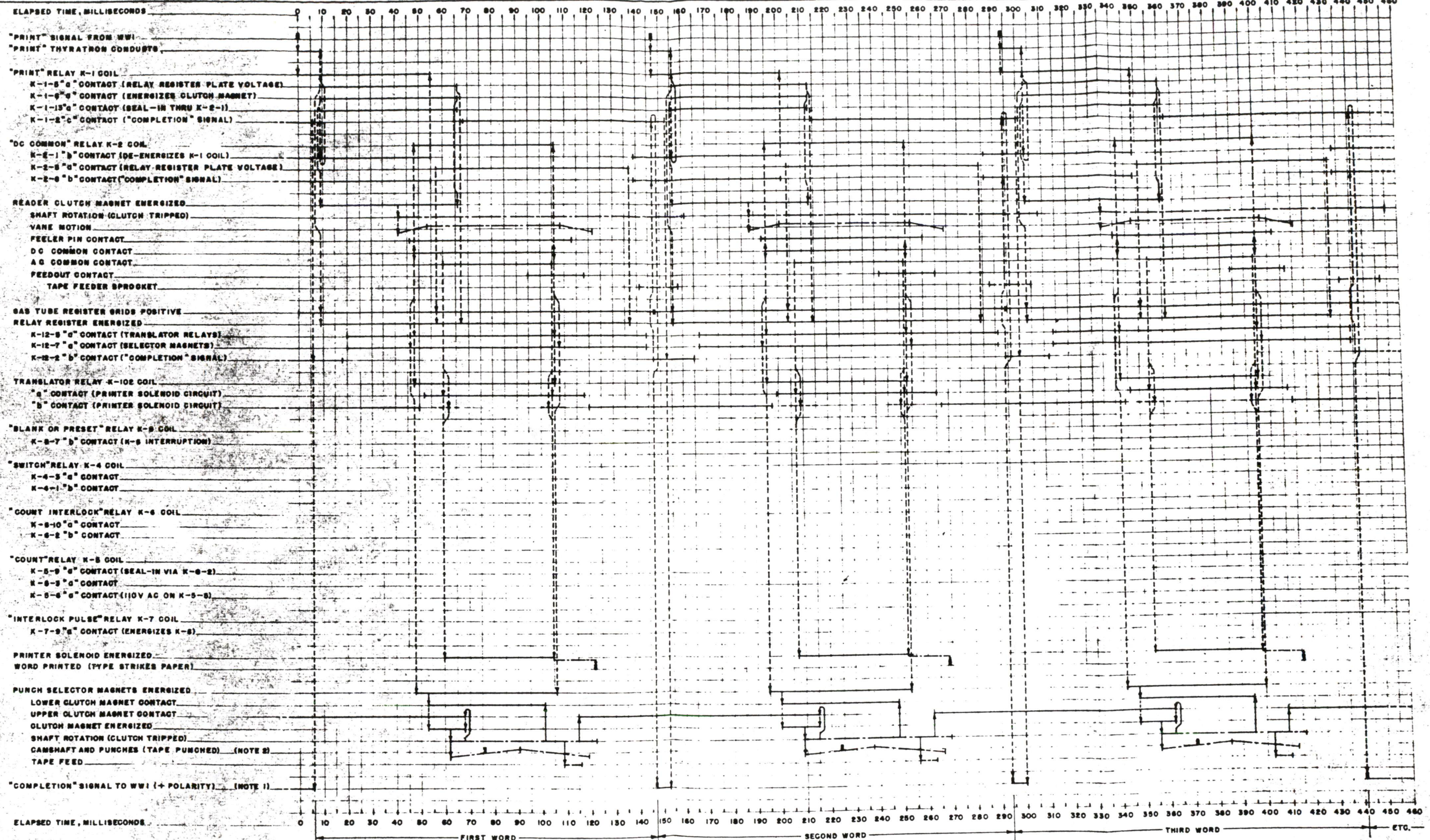
NOTE 2. SYMBOL (___L_._) DENOTES INSTANT AT WHICH PUNCH PERFORATES PAPER TAPE.

SYMBOLISM:

- DENOTES PERIOD DURING WHICH A RELAY COIL IS ENERGIZED, OR AN "a" OR "b" CONTACT CLOSED.
- OR / OR — VERTICAL DOTTED LINES INDICATE SIGNAL FLOW PATH OR AN OPERATION OF ONE CIRCUIT COMPONENT AS A RESULT OF OPERATION OF ANOTHER COMPONENT.
- DENOTES BLOCKING OF SIGNAL BY OPEN CONTACT.
- DENOTES SIGNAL THAT WOULD HAVE BEEN PRESENT HAD CONTACT BEEN CLOSED.
- COIL / "b" CONT / "a" CONT / "b" — COMPLETE INDICATION OF RELAY CONTACT OPERATING DELAY TIME WITH RESPECT TO RELAY COIL.
- DENOTES INITIATION OF ONE MECHANICAL FUNCTION AS RESULT OF ANOTHER BY MEANS OF HORIZONTAL DOT-DASH LINES AND VERTICAL SOLID LINES, AND ALSO MECHANICAL OPERATION OF CONTACTS BY EACH FUNCTION.
- K-1-13 "a" CONTACT. REFERS TO AN "a" CONTACT CONNECTED TO PIN 13 OF "PRINT" RELAY K-1, FOR EXAMPLE, AS SHOWN IN CIRCUIT SCHEMATIC R-35927.

CHANGED BY DATE	THIS IS A GRADED DRAWING OF HIGHEST GRADE APPROVED BELOW GRADE I FOR REFERENCE ONLY GRADE II PRELIMINARY DESIGN GRADE III FINAL DESIGN
SERVOMECHANISMS LABORATORY OF THE MASSACHUSETTS INSTITUTE OF TECHNOLOGY DIVISION OF INDUSTRIAL COOPERATION PROJECT NO. 6345	
TIMING DIAGRAM FOR "PUNCH WORDS & COMPLEMENTS, PRINT WORDS" MODE, TAPE OUTPUT EQUIPMENT, WWI	
SCALE —	A.M.G. 3-16-51
ENG. 4-4-51 J.L. Hanson	D-37303 B-REDUCTION

D-37308



NOTE 1. SYMBOL (L) SIGNIFIES INSTANT AT WHICH "COMPLETION" PULSE IS RETURNED TO WWI FROM PULSE GENERATOR (REF. DWG. R-35189). HORIZONTAL LINE (——) SIGNIFIES DURATION OF +90 VOLTAGE ON THIS CIRCUIT (TERMINAL E-1-12 ON JONES STRIP), -30V. BEING PRESENT AT ALL OTHER TIMES.

NOTE 2. SYMBOL (—L—) DENOTES INSTANT AT WHICH PUNCH PERFORATES PAPER TAPE.

SYMBOLISM:

DENOTES PERIOD DURING WHICH A RELAY COIL IS ENERGIZED, OR AN "a" OR "b" CONTACT CLOSED.

OR / OR — VERTICAL DOTTED LINES INDICATE SIGNAL FLOW PATH OR AN OPERATION OF ONE CIRCUIT COMPONENT AS A RESULT OF OPERATION OF ANOTHER COMPONENT.

DENOTES BLOCKING OF SIGNAL BY OPEN CONTACT.

DENOTES SIGNAL THAT WOULD HAVE BEEN PRESENT HAD CONTACT BEEN CLOSED.

COIL / "a" CONT / "b" CONT / "b" — COMPLETE INDICATION OF RELAY CONTACT OPERATING DELAY TIME WITH RESPECT TO RELAY COIL.

DENOTES INITIATION OF ONE MECHANICAL FUNCTION AS RESULT OF ANOTHER BY MEANS OF HORIZONTAL DOT-DASH LINES AND VERTICAL SOLID LINES, AND ALSO MECHANICAL OPERATION OF CONTACTS BY EACH FUNCTION.

K-1-13 "a" CONTACT. REFERS TO AN "a" CONTACT CONNECTED TO PIN 13 OF "PRINT" RELAY K-1, FOR EXAMPLE, AS SHOWN IN CIRCUIT SCHEMATIC R-35927.

GRADED BY DATE

THIS IS A GRADED DRAWING OF HIGHEST GRADE APPROVED BELOW
GRADE I FOR REFERENCE ONLY
GRADE II PRELIMINARY DESIGN
GRADE III FINAL DESIGN

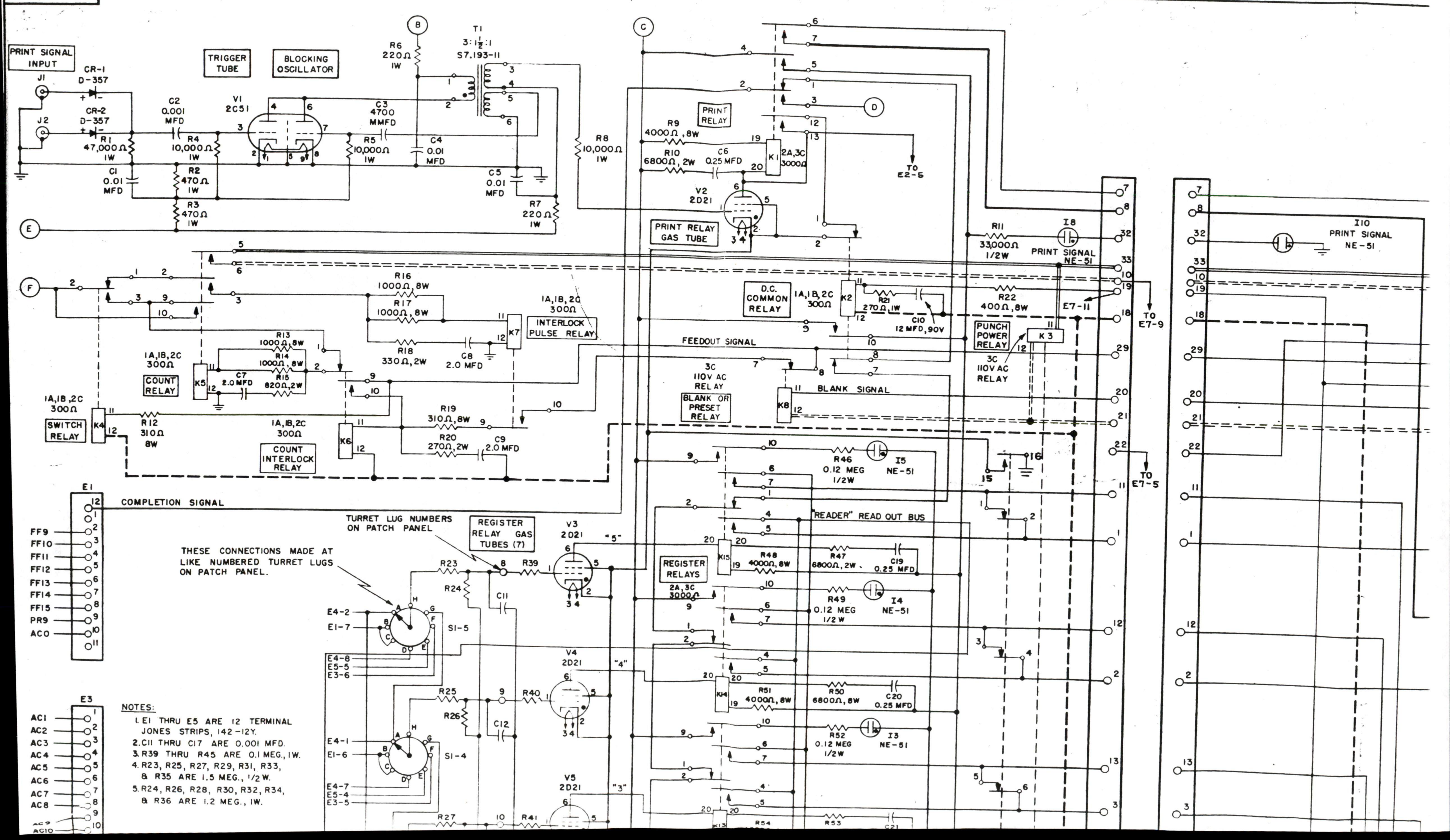
SERVOMECHANISMS LABORATORY OF THE
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
DIVISION OF INDUSTRIAL COOPERATION PROJECT NO. 6345

TIMING DIAGRAM FOR "WORDS ONLY" MODE, TAPE OUTPUT EQUIPMENT, WWI

SCALE —— DR. A.M.G. 3-23-51 ENG. 4-4-51 J.d. Hanson CK. APP.

D-37308
B-REDUCTION

R-35927-3

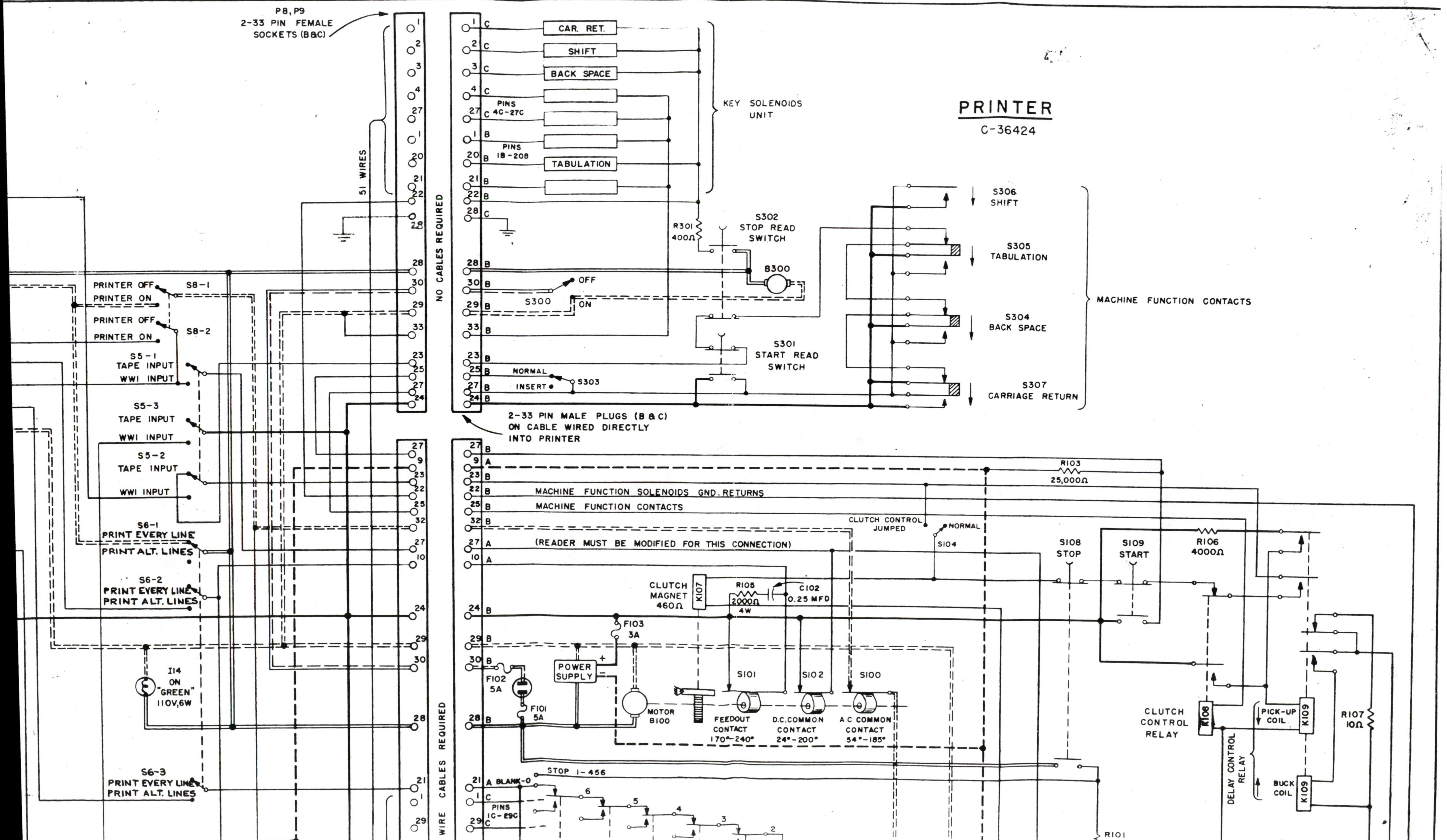


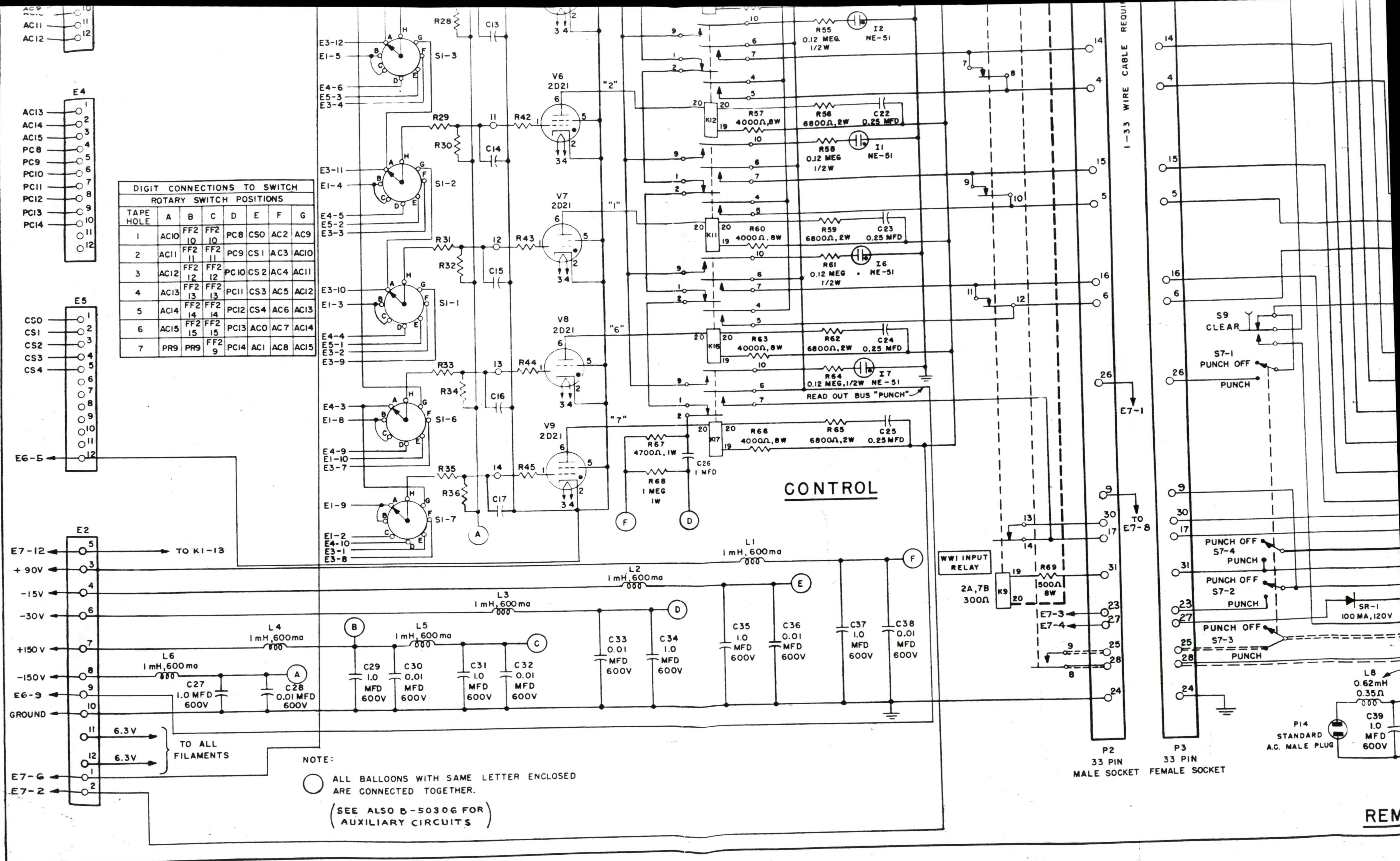
NOTES:

1. E1 THRU E5 ARE 12 TERMINAL JONES STRIPS, 142-12Y.
2. C11 THRU C17 ARE 0.001 MFD.
3. R39 THRU R45 ARE 0.1 MEG., 1W.
4. R23, R25, R27, R29, R31, R33, & R35 ARE 1.5 MEG., 1/2 W.
5. R24, R26, R28, R30, R32, R34, & R36 ARE 1.2 MEG., 1W.

PRINTER

C-36424



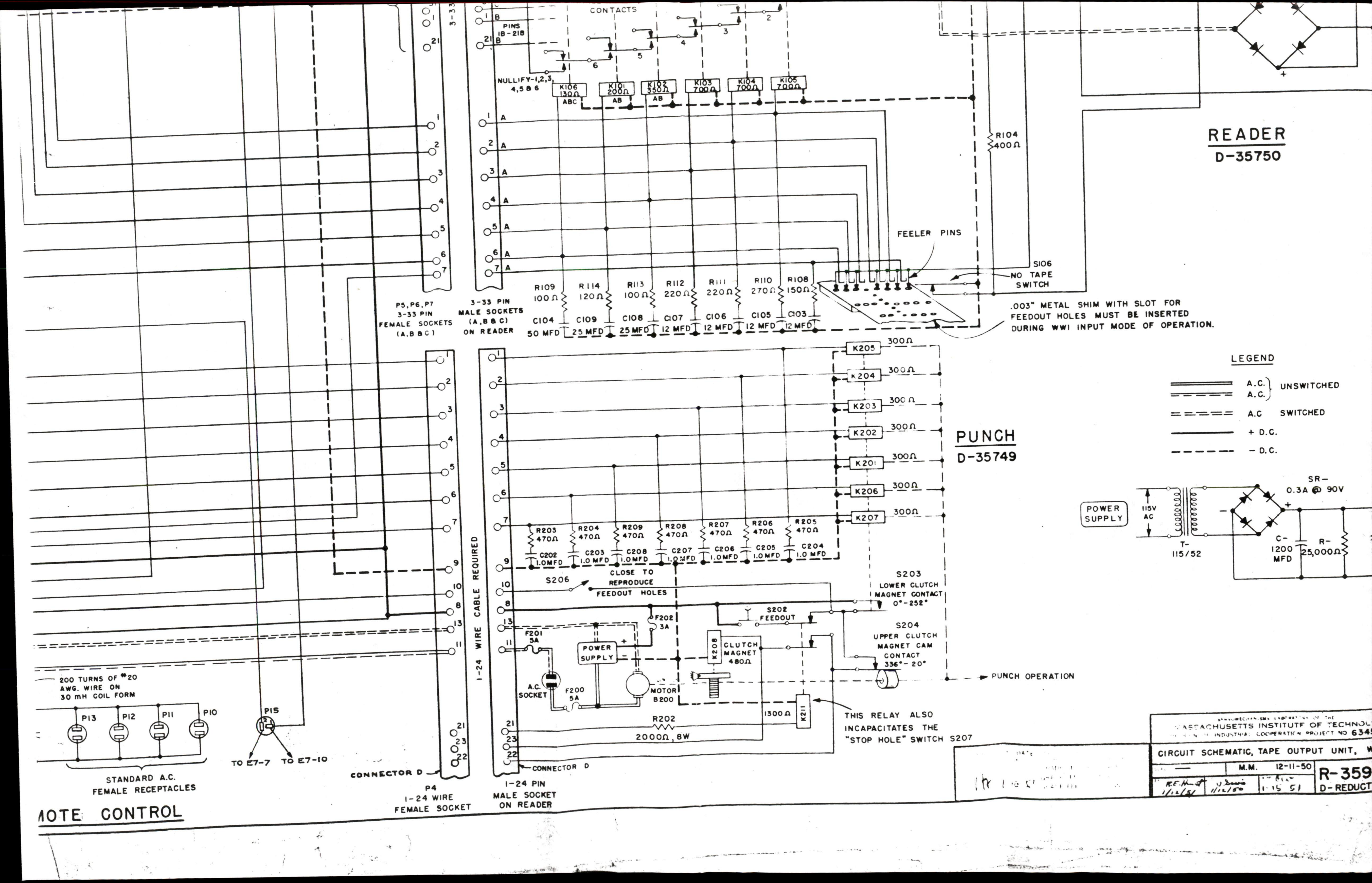


DIGIT CONNECTIONS TO SWITCH							
ROTARY SWITCH POSITIONS							
TAPE HOLE	A	B	C	D	E	F	G
1	AC10	FF2 10	FF2 10	PC8	CS0	AC2	AC9
2	AC11	FF2 11	FF2 11	PC9	CS1	AC3	AC10
3	AC12	FF2 12	FF2 12	PC10	CS2	AC4	AC11
4	AC13	FF2 13	FF2 13	PC11	CS3	AC5	AC12
5	AC14	FF2 14	FF2 14	PC12	CS4	AC6	AC13
6	AC15	FF2 15	FF2 15	PC13	AC0	AC7	AC14
7	PR9	PR9	FF2 9	PC14	AC1	AC8	AC15

NOTE:

◯ ALL BALLOONS WITH SAME LETTER ENCLOSED ARE CONNECTED TOGETHER.

(SEE ALSO B-50306 FOR AUXILIARY CIRCUITS)



CONTACTS
PINS 1B-21B
NULLIFY-1,2,3,4,5 & 6
K106 130Ω ABC
K101 200Ω AB
K102 350Ω AB
K103 700Ω
K104 700Ω
K105 700Ω
R104 400Ω
READER
D-35750
FEELER PINS
S106 NO TAPE SWITCH
R109 100Ω
R114 120Ω
R113 100Ω
R112 220Ω
R111 220Ω
R110 270Ω
R108 150Ω
C104 50 MFD
C109 25 MFD
C108 25 MFD
C107 12 MFD
C106 12 MFD
C105 12 MFD
C103 12 MFD
.003" METAL SHIM WITH SLOT FOR FEEDOUT HOLES MUST BE INSERTED DURING WWI INPUT MODE OF OPERATION.
P5, P6, P7 3-33 PIN FEMALE SOCKETS (A, B & C)
3-33 PIN MALE SOCKETS (A, B & C) ON READER
K205 300Ω
K204 300Ω
K203 300Ω
K202 300Ω
K201 300Ω
K206 300Ω
K207 300Ω
PUNCH
D-35749
LEGEND
A.C. A.C. UNSWITCHED
A.C SWITCHED
+ D.C.
- D.C.
POWER SUPPLY
115V AC
T-115/52
SR-0.3A @ 90V
C-1200 MFD
R-25,000Ω
R203 470Ω
R204 470Ω
R209 470Ω
R208 470Ω
R207 470Ω
R206 470Ω
R205 470Ω
C202 1.0MFD
C203 1.0 MFD
C208 1.0MFD
C207 1.0MFD
C206 1.0MFD
C205 1.0MFD
C204 1.0 MFD
1-24 WIRE CABLE REQUIRED
S206
CLOSE TO REPRODUCE FEEDOUT HOLES
S203 LOWER CLUTCH MAGNET CONTACT 0°-252°
S202 FEEDOUT
S204 UPPER CLUTCH MAGNET CAM CONTACT 336°-20°
F202 3A
F201 5A
F200 5A
POWER SUPPLY
A.C. SOCKET
MOTOR B200
K208 CLUTCH MAGNET 480Ω
K211 1300Ω
R202 2000Ω, 8W
PUNCH OPERATION
THIS RELAY ALSO INCAPACITATES THE "STOP HOLE" SWITCH S207
200 TURNS OF #20 AWG. WIRE ON 30 mH COIL FORM
P13 P12 P11 P10
STANDARD A.C. FEMALE RECEPTACLES
P15
TO E7-7
TO E7-10
CONNECTOR D
P4 1-24 WIRE FEMALE SOCKET
1-24 PIN MALE SOCKET ON READER
CONNECTOR D
MOTE CONTROL
MASSACHUSETTS INSTITUTE OF TECHNOL
INDUSTRIAL COOPERATION PROJECT NO 634
CIRCUIT SCHEMATIC, TAPE OUTPUT UNIT, W
M.M. 12-11-50
R-359
D-REDUCT