IFAC

INTERNATIONAL FEDERATION
OF AUTOMATIC CONTROL

WARSZAWA 1969

New Physical-Technical Principles for Components

Fourth Congress of the International Federation of Automatic Control
Warszawa 16–21 June 1969

TECHNICAL SESSION
36

NOT

Organized by
Naczelna Organizacja Techniczna w Polsce

INTERNATIONAL FEDERATION OF AUTOMATIC CONTROL

New Physical-Technical Principles for Components

TECHNICAL SESSION No 36

FOURTH CONGRESS OF THE INTERNATIONAL
FEDERATION OF AUTOMATIC CONTROL
WARSZAWA 16 – 21 JUNE 1969

Organized by
Naczelna Organizacja Techniczna w Polsce

Contents

Wydawnictwa Czasopism Technicznych NOT - Polska

Zakład Poligraficzny WCT NOT. Zam. 54/69.

СРАВНИТЕЛЬНЫЕ СВОЙСТВА, ОСНОВНЫЕ ХАРАКТЕРИСТИКИ И ОБЛАСТИ ПРИМЕНЕНИЯ ЭЛЕКТРОХИМИЧЕСКИХ ПРЕОБРАЗОВАТЕЛЕЙ

А.П.Шорыгин
Институт автоматики и телемеханики, Москва, СССР

Результатом расширения и усложнения задач автоматического управления и контроля явилось возникновение потребности в простых малогабаритных элементах для получения, преобразования и хранения информации для длительных непрерывных или дискретных процессов при низких уровнях полезных сигналов и (или) низких частотах. Выдвижение этих и некоторых других специфических требований к элементам, узлам и приборам, связанных с новыми задачами автоматического контроля и управления, привело к тому, что несмотря на интенсивное развитие вакуумной и полупроводниковой электроники, за последние годы проявился интерес к применению электрохимических явлений для построения преобразователей, удовлетворяющих современным условиям. В электрохимических элементах используются процессы в электролитах (жидких или твердых) и в полярных жидкостях I . Эти процессы могут быть полностью или частично обратимыми или практически необратимыми и сопровождаются изменением концентраций компонентов растворов, растворением и осаждением металлов, образованием и разрушением оксидных пленок, перемещением жидкостей и др. В настоящее время разрабатываются и частично применяются разнообразные электрохимические элементы для восприятия внешних воздействий, для усиления и преобразования сигналов, для хранения значений величин, для самонастройки, дистационного и телеуправления и др. Проводятся работы по созданию управляемых электрохимических матриц, делаются попытки использования электрохимических элементов для моделирования биологических объектов, ведутся поиски электрохимических систем и процессов наиболее пригодных для создания преобразователей.

Существующие электрохимические элементы можно подразделить на две группы. В первую группу входят элементы, имеющие выраженные преимущества или особые свойства, обеспечивающие их промышленное применение уже в настоящее время. Сюда можно отнести:

I) интегрирующие элементы с визуальным отсчетом или с фо-

тоэлектрическим или электрическим считыванием,

2) электрические выпрямительные диоды для очень малых токов низкой частоты,

3) датчики переменных или импульсных давлений и градиентов давлений низкой и инфранизкой частоты,

4) датчики состава кондуктометрические, гальванические и др.

Существенный интерес для ряда областей применения представляют, в первую очередь, электролитические интеграторные элементы:

$$P_{вых} = \varphi\left[\int_0^{\tau} i_{вх}(t)\,dt\right] \qquad (I)$$

где – выходной ток, – время его протекания.

Электрохимические интеграторные элементы существенно различаются по своим параметрам и практическим возможностям в зависимости от типа электрохимической системы, используемого эффекта и способа считывания (табл.I).

У газофазных интегральных диодов выходным параметром является смещение столбика электролита Δx в капилляре, соединяющем электродные камеры, заполненные водородом:

$$\Delta x \approx \frac{RT}{K_д p S_к} \int_0^{\tau} i_{вх}(t)\,dt \qquad (2)$$

где p – давление водорода, $S_к$ – площадь поперечного сечения капилляра, $K_д$ – константа прибора.

Характерные параметры газофазных элементов см.в табл.2. Такие элементы выпускаются промышленностью.

Окислительно-восстановительный процесс на рабочих электродах жидкофазных интеграторных элементов протекает по схеме

$$A_1 + ne \rightarrow A_2$$
$$A_2 - ne \rightarrow A_1 \qquad (3)$$

где A_1 – окисленная и A_2 – восстановленная форма компонентов раствора. Для типичных условий, когда скорость процесса определяется практически только диффузией к катоду формы A_1 концентрационная э.д.с. на выходе интеграторного диода определяется по уравнению:

$$C_{01} n F S \ell \cdot th \frac{E_c}{E_T} = \int_0^{\tau} i_{вх}(t)\,dt \quad (4)$$

причем $E_T = \frac{RT}{nF}$, F - число фарадея, S и ℓ - площадь поперечного сечения и длина камер, разграниченных пористой перегородкой, C_{01} - начальная концентрация формы A_1 (лесловного носителя). При $E_c << E_T$ зависимость становится линейной и чувствительность интеграторного диода:

$$S_E = \frac{E_c}{q_{вх}} = \frac{2RT}{C_{01}(nF)^2 S \ell} \quad (5)$$

Максимальное значение входного заряда определяется объемом камер $V = S\ell$ и исходной концентрацией C_{01}

$$(q_{вх})_{max} = nFC_{01}V \quad (6)$$

Для обеспечения необходимой точности интегрирования верхний предел спектра частот входного сигнала должен удовлетворять условию:

$$f_2 << \frac{D}{2\pi\ell^2} \quad (7)$$

(D - коэффициент диффузии).

Время хранения значения интеграла определяется постоянной времени саморазряда:

$$\tau_{CR3} = \frac{V l_к}{D S_к} \qquad (8)$$

Здесь $S_к$ - суммарная эквивалентная площадь поперечного сечения капилляров перегородки, разделяющей катодную и анодную камеры, $l_к$ - толщина перегородки.

Выходной параметр жидкофазных триодов и тетродов - ток диффузии $I_д$ между анодом интеграторного элемента и считывающим электродом в анодной камере, на который подается отрицательное смещение $E_т$. Его характеристика:

$$(I_д)_{вых} = \frac{2 q_{вх} D}{l_{CA}^2} \cdot \frac{e^{U_c/E_т} - e^{U_A/E_т}}{e^{U_c/E_т} + e^{U_A/E_т}} \qquad (9)$$

где U_c и U_A напряжения электрод-электролит для считывающего электрода и для анода, а l_{CA} - расстояние между этими электродами. Так как ток в выходной цепи ограничен сопротивлением нагрузки $R_н$ то при достижении входным зарядом значения

$$q_{вх} \geqslant \frac{E}{R_н} \cdot \frac{l_{CA}^2}{2D} \qquad (10)$$

наступает насыщение и предельный ток диффузии:

$$I_{дп} = \frac{n F D C_{ан} S_{сл}}{\ell_{сл}} \qquad (11)$$

В этих условиях, практически используемых при интегрировании:

$$I_{вых} = S_I \int_0^{\tau} i_{вх}(t)\,dt \qquad (12)$$

($S_I = \frac{k_т n D}{\ell_{сл}}$, а $k_т$ - константа элемента). Характерные значения основных параметров триодов и тетродов - см.табл.2.

Твердофазные интеграторные элементы представлены в настоящее время наибольшим разнообразием типов (табл.I, 2). Окислительно-восстановительный процесс протекает у них по схеме:

$$Me - ne \rightarrow Me^{n+}$$
$$Me^{n+} + ne \rightarrow Me \qquad (13)$$

Масса вещества, растворенного или осажденного на электроде за время τ (в частности за время одного цикла у диода с дискретным считыванием)

$$m_{ме} = \frac{A\eta}{nF} \int_0^{\tau} i_{вх}(t)\,dt \qquad (14)$$

где A - атомный вес, η - коэффициент выхода по току
В случае капиллярной ячейки, выходной параметр - смещение границы раздела электрод-раствор:

$$\Delta x_{вых} = S_\ell \int_0^{\tau} i_{вх}(t)\,dt \qquad (15)$$

причем $S_\ell = \frac{A\eta}{nF\gamma S_к}$ - чувствительность (здесь γ - плотность столбика вещества, $S_к$ - площадь его поперечного сечения). При использовании такого элемента в качестве счетчика времени наработки:

$$t_{нар} \approx \frac{\Delta x_{вых}}{S_\ell \frac{V_{ст}}{R_д} \frac{1}{1 + R_я / R_д}} \qquad (16)$$

где $U_{ст}$ – напряжение сети, $R_з$ – ограничивающее сопротивление, $R_я$ – сопротивление ячейки.

Особый интерес представляют твердофазные канальные триоды с монотонным изменением линейного сопротивления $R_{сч}$ на выходе в зависимости от электрического заряда, протекшего во входной цепи [2, 3]. При достаточно малом $R_{сч}$ и равномерном осаждении металла на электроде хранения:

$$\sigma_{сч} = \frac{R_{сч}(t)}{R_{сч}(0)} = \frac{1}{1 + \frac{R_{сч}(0)}{\psi}\int_0^{\tau} i_y(t)\,dt} \qquad (17)$$

где $R_{сч}(0) = (R_{сч})_{max}$ – значение выходного сопротивления в начальный момент, $R_{сч}(t)$ – то же в момент t, а

$$\psi = \frac{nF}{nA}\cdot\rho\gamma l_x^2$$ (причем γ – средняя плотность осажденного металла, ρ – его удельное сопротивление l_x – длина электрода хранения). При достаточно больших q_{ox} скорость изменения сопротивления [3]:

$$\frac{d\sigma_{сч}}{dt} = \frac{\psi}{R_{сч}(0)\cdot I_{y=}\cdot t^2} \qquad (18)$$

и наименьшее время полного изменения сопротивления:

$$t_{min} = \frac{\psi\left[\frac{R_{сч}(0)}{(R_{сч})_{min}} - 1\right]}{R_{сч}(0)\cdot(I_{y=})_{max}} \qquad (19)$$

($(I_{y=})_{max}$ – наибольший допустимый ток управления).

При типичных условиях сопротивление по цепи считывания для переменного тока [4]:

$$Z_{сч} \approx 2\,\frac{\sqrt{R_{x1}A_{\phi 1}}}{f^{0,25}}\cdot th\left(\frac{l_x}{2}\sqrt{\frac{R_{x1}}{A_{\phi 1}}}\cdot f^{0,25}\right) \qquad (20)$$

где l_x – длина электрода хранения (резистивного электрода), R_{x1} – сопротивление его единицы длины, $A_{\varphi 1}$ – коэффициент, характеризующий фарадеевский импеданс электрод-электролит на единицу длины электрода.

По диапазону времен интегрирования, величинам токов, точности, габаритам электролитические интеграторные элементы в ряде случаев имеют явные преимущества (см.рис.I). Они обладают высокой линейностью, очень малым собственным потреблением (десятки или сотни микроватт), пригодностью для интегрирования долей микрокулонов, могут работать от вентильных фотоэлементов, термопар, датчиков Холла и др. Габаритные размеры их не превышают размеров миниатюрных электронных ламп, а вес составляет 2 ÷ 20 г. Наибольшая точность достигается у твердофазных интеграторных элементов с дискретным считыванием.

Процессы в электрохимических элементах протекают в весьма тонких слоях, измеряемых долями микрона, так что можно ожидать появления значительно более миниатюрных элементов ("планарных").

Широкое применение получат в ближайшее время основанные на рассмотренных принципах счетчики машинного времени (в связи с необходимостью контролировать, в соответствии с ГОСТами и нормалями, работоспособность и срок службы различных приборов, аппаратов и машин) (рис.2) [6], а также интегрирующие звенья, особенно с большими временами интегрирования, узлы настройки адаптивных систем и оптимизаторов на основе твердофазных канальных триодов, узлы задержки и др. Существенный интерес представляют также электрохимические интеграторные устройства для автоматической обработки измерительной информации, реализующие алгоритм вида:

$$y_i = \frac{k_i \int_{t_i}^{t_{i+1}} [x(t) - x(t_0)]\,dt}{\sum_{1}^{n} k_j \int_{t_j}^{t_{j+1}} [x(t) - x(t_0)]\,dt} \qquad (21)$$

В жидкофазных электрохимических выпрямительных диодах используется окислительно-восстановительный процесс по схеме (2). Если скорость процесса определяется практически только диффузией, то интегральное уравнение для нахождения тока через диод [5]:

$$\left[1-\frac{1}{U_1}\cdot\frac{d}{dt}\int_0^t f_1(t-\tau)\cdot i(\tau)\,d\tau\right]\cdot\exp\frac{nFU_{вх}}{RT}=$$

$$=\left[1-\frac{1}{U_2}\cdot\frac{d}{dt}\int_0^t f_2(t-\tau)\cdot i(\tau)\,d\tau\right]^3, \qquad (22)$$

где

$$U_{1,2}=nFSC_{1,2}\;;\; f_{1,2}=\frac{a}{D_{1,2}}\left\{1-\exp\frac{tD_{1,2}}{a^2}\cdot\operatorname{erfc}\sqrt{\frac{tD_{1,2}}{a^2}}\right\} \qquad (23)$$

причем a - радиус сферического микроэлектрода, а индексами I и 2 отмечены величины, относящиеся соответственно к формам A_1 и A_2

При $t\to\infty$, $\lim i(t)=i_\infty$ и

$$\lim\frac{1}{U_{1,2}}\cdot\frac{d}{dt}\int_0^t f_{1,2}(t-\tau)i(\tau)\,d\tau=\frac{a\,i_\infty}{nFC_{1,2}D_{1,2}} \qquad (24)$$

так что статическая вольтамперная характеристика диода

$$E_\infty=E_T\left[3\ln\left(1-\frac{a\cdot i_\infty}{nFSC_{02}D_2}\right)-\ln\left(1-\frac{a\cdot i_\infty}{nFSC_{01}D_1}\right)\right] \qquad (25)$$

Электролитические выпрямительные диоды позволяют выпрямлять весьма малые токи - от сотых или десятых долей микроампера до сотен микроампер, что обеспечивает им существенное преимущество (при низких частотах), по сравнению с диодами, построенными на других принципах. На основе электролитических диодов этого типа созданы элементы, реализующие операции $\frac{1}{\sqrt{p}}x$ и $\sqrt{p}\,x$ в широком диапазоне инфранизких частот [5], а также элементы с отрицательным сопротивлением.

Электролитические и электрокинетические нерезонансные датчики малых переменных и импульсных давлений, датчики градиентов давлений (в частности электроакустические преобразователи инфранизкой и звуковой частот, пригодные для работы при очень высоких статических давлениях) и датчики ускорений, как видно из рис.3, 4, 5, имеют преимущества по ширине возможного частотного диапазона и охвату низких частот, хотя и уступают пьезорезистивным преобразователям по чувствительности. Характеристика вход-выход для электрокинетических преобразователей линейна в широком диапазоне амплитуд. Для холостого хода:

$$E_{вых} = \frac{K_{экп}\,\xi\,\varepsilon\,\Delta p_{вх}}{4\pi\mu\lambda_\sigma} \qquad (26)$$

где ξ - электрокинетический потенциал, ε, μ, λ_σ - диэлектрический коэффициент, динамическая вязкость и результирующая электропроводность рабочей жидкости, $K_{экп}$ - коэффициент, учитывающий влияние гидродинамических свойств перегородки. Амплитудно-частотная характеристика этого перобразователя:

$$E_m = \frac{\xi\varepsilon}{4\pi\mu\lambda_\sigma} \cdot \frac{\omega\tau\rho_m}{\sqrt{1+(\omega\tau_м)^2}} \qquad (27)$$

($\tau_м$ - механическая постоянная времени перобразователя).

Характеристика вход-выход жидкофазных электролитических датчиков (работающих по схеме (2)) для одной из конфигураций катодного канала:

$$I_{вых} = \frac{2FC_0\delta^3 в}{3\mu a\pi^3}\cdot\Delta p_{ex}\left[1-\exp\left(-\frac{12\pi^2 D\mu a^2}{\delta^4}\cdot\frac{1}{\Delta p_{вх}}\right)\right] \qquad (28)$$

где C_0 - концентрация неосновных носителей в свободном объеме, D - коэффициент диффузии, δ, a, $в$ - геометрические размеры щелевого канала, μ - динамическая вязкость раствора. Изменяя форму катодного канала можно получать различные амплитудные характеристики (линейные, логарифмические и др.).

Указанные элементы в различном исполнении получают применение в качестве сейсмодатчиков, датчиков биологических параметров, акустических приемников инфразвуковой частоты и др., из которых электрокинетические - обладают высокой линейностью амплитудных характеристик и наиболее широким динамическим диапазоном, а электролитические - очень низким частотным диапазоном. Однако и те и другие в настоящее время уступают новым образцам пьезоэлектрических и пьезорезистивных датчиков по рабочему диапазону температур.

Широко применяются и выпускаются в лабораторном и промышленном исполнении различные типы электрохимических концентрато-

меров для растворов и газов (переменного тока, контактные и бесконтактные, а также гальванические).

Вторую группу образуют электрохимические элементы, требующие дополнительного изучения и улучшения их характеристик. Можно ожидать, что соответствующие исследования позволят определить такие области применения, кроме указанных выше, где окажется целесообразным использовать электрохимические преобразователи. В эту группу можно, например, отнести:

1) электролитические датчики вибраций;
2) электрокинетические датчики скорости;
3) индикаторы малых напряжений;
4) инфранизкочастотные жидкостные усилительные элементы;
5) электролитические и электрокинетические функциональные элементы;
6) твердофазные электролитические исполнительные элементы и усилители мощности и др.

Для развития данной области автоматики необходимо проведение исследований в следующих основных направлениях:

а) изыскание новых принципов построения электрохимических систем для расширения технических возможностей электрохимических элементов;

б) исследование общих свойств и характеристик датчиков, преобразователей и других элементов и нахождение наиболее целесообразных областей их применения;

в) разработка теории и методов расчета электрохимических элементов различных принципов действия;

г) разработка схем с использованием электрохимических элементов в устройствах автоматического управления и контроля и методов расчета таких схем;

д) разработка новейшей технологии изготовления электрохимических элементов, обеспечивающей их высокую надежность.

Заключение

Проведенное сравнение характеристик дает возможность заключить, что на основе электрохимических систем возможно строить разнообразные датчики и преобразователи информации для процессов с очень низкими частотами (от десятых или сотых долей герца и менее), с очень малыми токами (до долей микро-

ампера), с очень большими временами (до десятков или сотен суток) при сохранении весьма малых габаритов и малом потреблении энергии. Вместе с тем такие системы позволяют создавать элементы с управлением и неразрушающим считыванием микросекундными импульсами.

Цитированная литература

1. А.П.Горшгин. Электрохимические элементы (основные особенности; классификация), ЭИКА вып.8, изд.Энергия, 1967.
2. А.П.Горшгин. Электрохимические элементы. АиТ, № 8, 1966.
3. В.С.Боровков, В.В.Трейер. Электрохимические аналоговые запоминающие устройства (редактор А.П.Горшгин). Изд. ВЗЭИ, М., 1967.
4. А.П.Горшгин. Исследование входных, выходных и передаточных характеристик твердофазных капельных электролитических триодов. Электрохимия, т.IV, 1968 №12.
5. Р.Ш.Нигматуллин. Теория электрохимического диода. ДАН, т.150, 603, 1963, № 3.
6. А.П.Горшгин, З.В.Казарян. Счетчики машинного времени. Труды АН Арм.ССР (принято к печати).
7. Х.С.Русейнов, М.С.Касимзаде. Переходные процессы в электрокинетических сейсмоприемниках. Автометрия, стр.19, 1966, № 4.

Таблица № I

Типы интеграторного элемента	Эффект, используемый для преобразования	Выходная физическая величина	Считывание
	Окислительно-восстановительные системы с инертными электродами		
I. Газофазные	а/ образование, выделение и поглощение водорода при электрохимических реакциях	а/ изменение разности объемов /или давлений/ газа над разноименными электродами	Визуальное
		б/ замыкание и размыкание электролитического контакта.	Электрическое
	б/ изменение свойств электрода при поглощении водорода, образующегося при электрохимических процессах	а/ изменение электропроводности электрода	Электрическое
		б/ изменение концентрационной э.д.с.	Электрическое
2. Жидкофазные	Изменение распределения концентрации компонентов электролита по объему электрохимической ячейки	а/ изменение концентрационной э.д.с.	Электрическое
		б/ изменение оптической плотности электролита	Визуальное или электрическое
		в/ изменение предельного тока диффузии	Электрическое

	Окислительно-восстановительные системы с электрохимически растворимыми электродами		
3. Твердофазные	а/ изменение массы вещества электрода за счет переноса	а/ изменение электропроводности электрода	Электрическое
		б/ изменение оптической плотности тонкопленочного электрода	Фотоэлектрическое
		в/ изменение линейных размеров электрода	Визуальное, Фотоэлектрическое, электрическое
		г/ изменение веса электрода	-
		д/ изменение электропроводности ячейки /за счет изменения геометрических размеров электрода/	Электрическое
	б/ изменение равновесного потенциала при анодном растворении анодно-активного металла на инертной подложке	Изменение падения напряжения в электрохимической ячейке	Электрическое

Таблица 2

Типы электрохимических интеграторных элементов	Максимальное достигнутое время интегр. (сек)	Основная погрешность (%)	а)Температурная погрешность (%/°С) б)Рабочий диапазон температур (°С)	Входной ток (а)	Наибольший входной заряд (к)
I	2	3	4	5	6
I.Газофазные интеграторные диоды (с визуальным отсчетом) и триоды (с разрушающим дискретным электрическим считыванием)	$10^5 \div 10^6$	2÷4	а) 0,I б)I5÷25	$10^{-5} \div 10^{-3}$	0,I÷I
2.Жидкофазные интеграторные диоды (с неразрушающим считыванием по концентрационной э.д.с.)и триоды (с неразрушающим считыванием по предельному току диффузии)	$10^6 \div 10^7$	3÷5	а)I,5÷2,3 б)0÷60 (с обычным электролитом)	$10^{-9} \div 10^{-4}$	0,05÷2
3.Жидкофазные интеграторные тетроды (с неразрушающим считыванием по предельному току диффузии)	$10^3 \div 10^4$	2÷3	тоже	$10^{-6} \div 10^{-4}$	0,I÷I
4.Твердофазные интеграторные диоды (с визуальным отсчетом)	$10^7 \div 10^8$	5÷I5	-	$10^{-5} \div 10^{-2}$	0,I÷I0
5.Твердофазные интеграторные диоды непрерывным неразрушающим электрическим и фотоэлектрическим считыванием)	$10^4 \div 10^8$	2÷5	а) 0,I б)-20÷60	$10^{-5} \div 10^{-2}$	0,I÷I0

I	2	3	4	5	6
6.Твердофазные интегральные диоды (дискретным разрушающим электрическим считыванием)	$10^4 \div 10^7$	0,01÷0,1	а)0,01÷0,1 б)-55÷75	$10^{-8} \div 10^{-4}$	0,4
7.Твердофазные диоды (с непрерывным неразрушающим фотоэлектрическим считыванием по оптической плотности электрода хранения)	$10^4 \div 10^5$	5÷10	а) 0,5 б) 0÷60 (с обычным электролитом)	$10^{-5} \div 10^{-2}$	1,0
8. Твердофазные канальные триоды (с непрерывным неразрушающим считыванием по сопротивлению электрода хранения)	$10^4 \div 10^5$	10÷15	то же	$10^{-5} \div 10^{-2}$	1,0

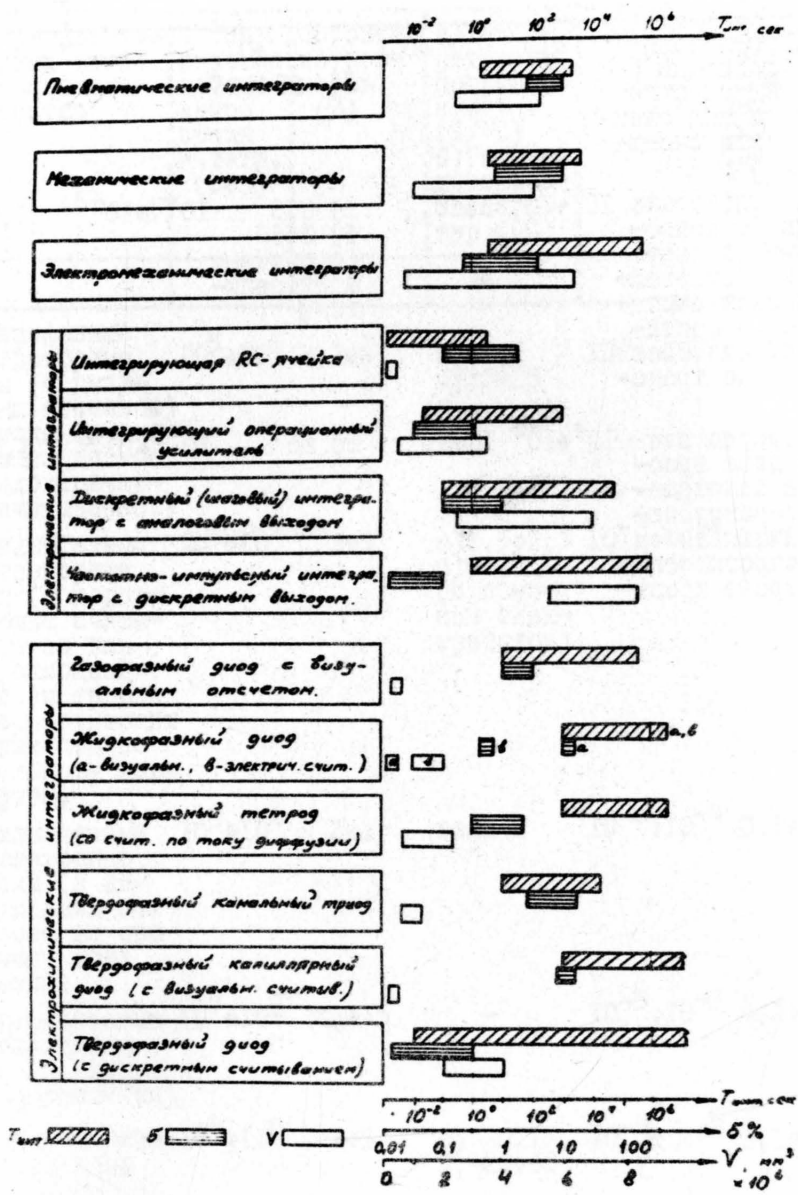


Рис.I Сравнение харак терных областей значений параметров интеграторных элементов, основанных на различных физических принципах.

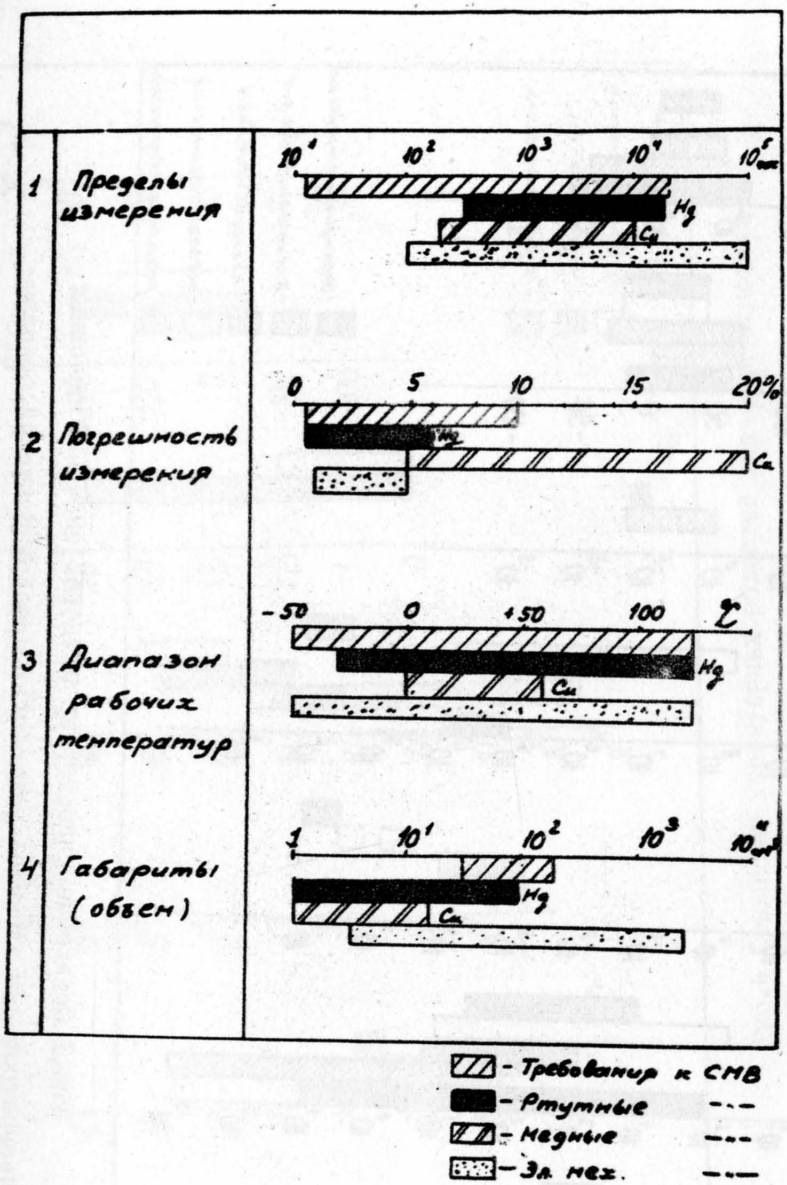


Рис.2 Сравнение по основным параметрам электрохимических и электромеханических счетчиков машинного времени, а также сопоставление этих данных с требованиями к таким счетчикам.

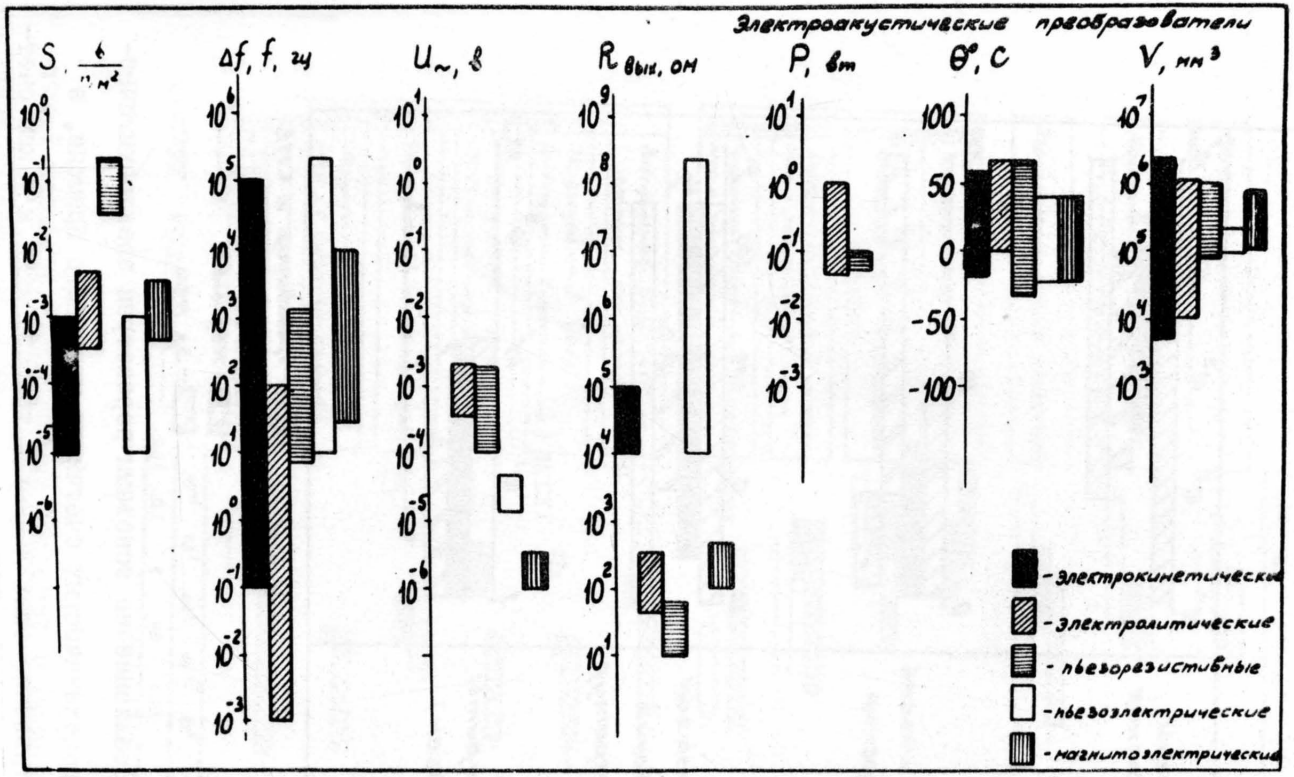


Рис.3 Сравнение характерных областей значений основных параметров для электроакустических преобразователей, основанных на различных физических принципах.

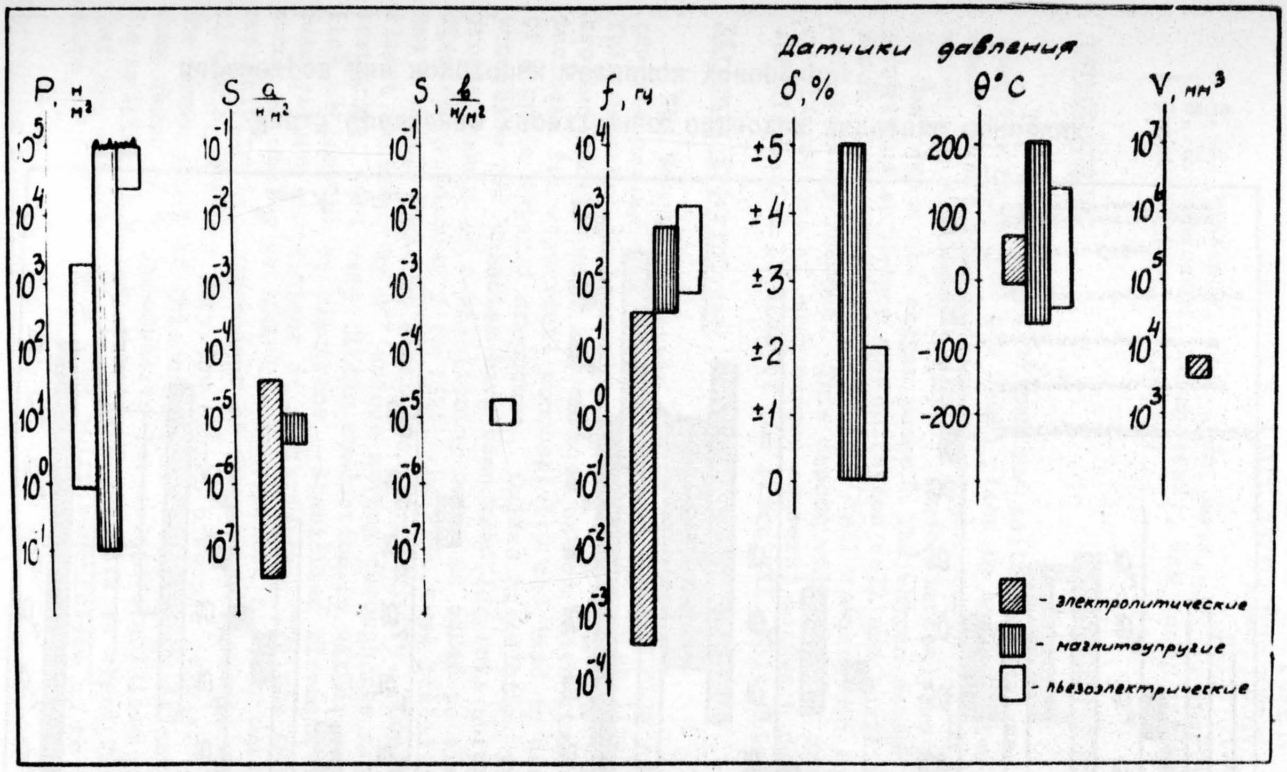


Рис.4 Сравнение характерных областей значений основных параметров для датчиков малых давлений, основанных на различных физических принципах.

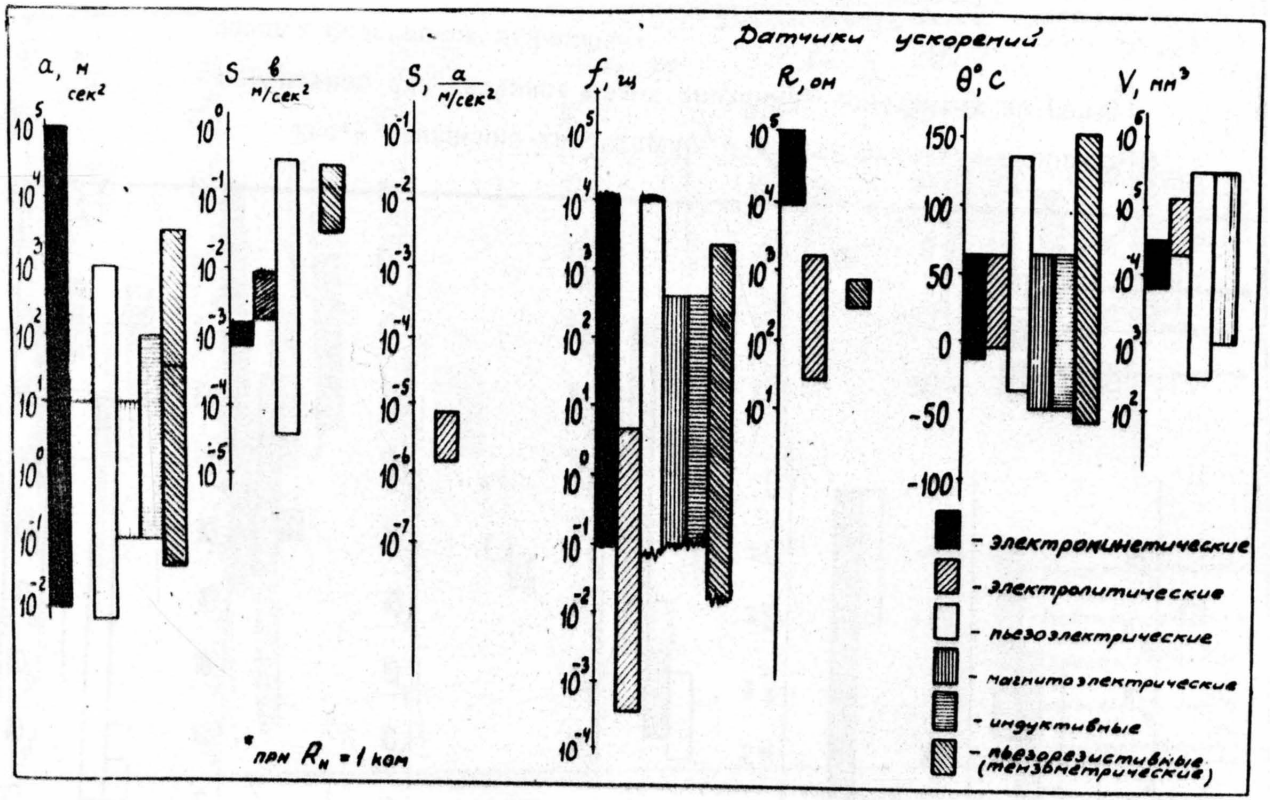


Рис.5 Сравнение характерных областей значений основных параметров для некоторых датчиков ускорений.

THE SOLID ELECTROLYTE OXYGEN SENSOR THEORY AND APPLICATION

S. J. Lawrence, West Lynn, Mass., H. S. Spacil and D. L. Schroeder, Schenectady, New York, all of the General Electric Company

INTRODUCTION

The development of various solid electrolytes which are specific to oxygen has taken place in the past two decades. These permit construction of oxygen concentration cells, which are measuring devices. These electrolytes are also used in fuel cells and oxygen pumps. The oxygen sensor is a device comprising an oxygen concentration cell and auxiliary equipment which includes sample handling, cell temperature control and cell voltage readout. This report will present the theoretical considerations of such oxygen sensors[1], hardware description and discussion of some of the practical applications.

Zirconia Electrolytes

The General Electric Oxygen Sensor contains a solid electrolyte oxygen concentration cell using "stabilized" zirconia. This material is basically zirconium dioxide, ZrO_2, with controlled additions of one or more oxides which simultaneously increase its electrical conductivity and lead to development of a cubic crystal structure. Oxides of calcium, yttrium, ytterbium and scandium have been used for this purpose. If the proper amount of a stabilizing oxide is added to zirconium dioxide (typically 5 to 15%), the resulting stabilized zirconia has a unique type of electrical conductivity. Conduction is almost entirely by O^{--} anions moving through the crystal structure, with very little conduction by either positive (metallic) cations or by electrons. Stabilized zirconia has a negative temperature coefficient of resistivity. Typical resistivity is about 10 ohm-cm at 1000°C. In contrast to many oxides, the resistivity is independent of the oxygen pressure over the zirconia, down to very low oxygen pressure levels. At 1000°C, the lower limit of this oxygen pressure range is not higher than about 10^{-21} atmospheres for calcia stabilized zirconia[2]. The melting point of stabilized zirconia is about 2200°C, so that the material is a true solid electrolyte up to this temperature.

Operating Principle

The high oxygen ion conductivity of stabilized zirconia allows it to be used for measuring oxygen pressures or con-

centrations in gas mixtures. Figure 1 is a schematic cross section of an oxygen concentration cell consisting of a stabilized zirconia electrolyte with electrodes attached to two opposite surfaces which are exposed to different oxygen pressures. The closed-end tube configuration is convenient; the inner electrode is exposed to a gas with known oxygen pressure, such as air. It then serves as a reference electrode. The potential of the electrode exposed to the sample gas is measured with respect to the reference electrode. In the cell of Figure 1, the open circuit EMF is given by

$$E_o = \frac{CRT}{4F} \log \frac{[O_2]}{[O_2]_R} \qquad (1)$$

where

$[O_2]$ = oxygen pressure of sample gas, atmospheres
$[O_2]_R$ = reference oxygen pressure, atmospheres
C = 2.3026 (conversion from base 10 to natural logarithm)
R = 8.313 (gas constant in joules/mol)
T = cell temperature ^oK
F = 96501 (Faraday constant in coulombs/equivalent)

For the equation to apply, the following conditions must be satisfied:

1. Oxygen ion conductivity must predominate in the electrolyte.

2. The electrolyte must be leak tight since any leaks would allow oxygen to transfer through the electrolyte without forming ions.

3. Reversibility of the electrode to allow gaseous oxygen to ionize freely at the electrolyte surface.

4. Low cell resistance to permit voltage measurements with existing equipment.

The equation may be rearranged to express the oxygen pressure in the sample gas:

$$\log [O_2] = \log [O_2]_R + \frac{4F}{CRT} E_o \qquad (2)$$

Thus, if the reference oxygen pressure and the temperature are known, measurement of the open circuit voltage can be used to determine the oxygen pressure in the sample gas. A negative voltage indicates that the sample oxygen pressure is less than the reference oxygen pressure.

Operating Conditions

A temperature of 850°C generally allows satisfactory operation of this concentration cell, and does not present serious material problems. Porous platinum electrodes are satisfactory. Reference oxygen may be supplied by flow of a gas of known oxygen content (for example, air) or a metal/metal oxide mixture that serves as both a metallic conductor and a source of oxygen due to the dissociation of the oxide. At 850°C, the coefficient of the log term in Equation (1) is about .056 volts; thus changing the oxygen pressure in the cell by a factor of 10 changes E_o by about 56 millivolts. The magnitude of the cell voltage depends on the reference gas used. With air reference, at .21 atmospheres oxygen partial pressures, measurement of the equivalent of oxygen pressures as low as 10^{-24} atmospheres is allowable, resulting in an open circuit voltage of about 1300 millivolts. A nickel/nickel oxide reference at 850° generates an oxygen pressure of about 10^{-13} atmospheres, and allows measurement to 10^{-27} equivalent atmospheres with a corresponding open circuit voltage of about 800 millivolts. The term equivalent is used because these low pressures are measured in reactive systems (which will be explained later) where oxygen exists in equilibrium with water-hydrogen or carbon monoxide-carbon dioxide. Where oxygen is diluted in a relatively inert gas such as nitrogen, 10^{-7} atmospheres can be measured. Figure 2 illustrates the range of equilibrium O_2 partial pressure at 1000°C in several commonly encountered industrial atmospheres[3]. The continuous measurement of O_2 partial pressure in these gases will develop better knowledge about the processes involved and lead to more effective control.

Hardware

Figure 3 is a photograph of two experimental cells. Both are designed to fit inside heaters which bring the cell temperature to the operating level. One is a closed end type; the other is a flow-through type.

For the most part, we have found that the tube construction for the electrolyte was desirable. The tubular shape provides strength and rigidity for brittle materials such as ZrO_2; it also minimizes the amount of joining and sealing required. We found that commercially available zirconia tubes often leaked beyond the rigid specification we have set. Therefore, we have developed our own process for making these.

The process is based on plasma arc spraying onto mandrels, which are subsequently leached out, followed by sintering of the free-standing body to render the zirconia electrolyte leak tight[4]. Commercial grade powders are spray-dried to prepare free material for the plasma-arc spray operation. Stabilizing

and sintering agents are added.

O_2 Partial Pressure and O_2 Content

This oxygen sensor measures the equilibrium O_2 partial pressure of the sample gas at the temperature and pressure of the cell. Any composition not at equilibrium will rearrange under the cell conditions. Thus, oxygen in the system $H_2O/H_2/O_2$ or $CO/CO_2/O_2$ or similar, will redistribute according to the equilibrium constant for the system at operating conditions. In a gas consisting of oxygen and an inert diluent such as nitrogen:

$$[O_2] = xP \tag{3}$$

where x = mol fraction oxygen in the gas

P = total gas pressure

In the water-hydrogen-oxygen system, the equilibrium oxygen pressure is:

$$[O_2] = \frac{1}{K_1^2} \frac{[H_2O]^2}{[H_2]^2} \tag{4}$$

where K_1 = equilibrium constant for the reaction:

$$H_2 + 1/2O_2 \rightleftharpoons H_2O \tag{I}$$

In the carbon monoxide-carbon dioxide-oxygen system, the equilibrium oxygen pressure is:

$$O_2 = \frac{1}{K_2^2} \frac{[CO_2]^2}{[CO]^2} \tag{5}$$

where K_2 = equilibrium constant for the reaction:

$$CO + 1/2O_2 \rightleftharpoons CO_2 \tag{II}$$

Logarithmic Units for O_2 Partial Pressure

One way of expressing the oxygen partial pressure is smaller to the pH value of aqueous chemistry, and is defined as:

$$pO = \log \frac{1}{[O_2]^{\frac{1}{2}}} \tag{6}$$

For most cases of interest, the pO value is positive, dimensionless, and is in the range 0-14.

Another logarithmic expression with units of calories per mol is defined as:

$$e0 = 4.575T \log ([O_2]^{\frac{1}{2}}) \qquad (7)$$

The e0 value will normally be negative and in the range 0 to -100 kcal/mol.

One reason for the use of these expressions is the general reaction expressing the oxidation of a metal by O_2:

$$(a/b)M + (1/2)O_2 \longrightarrow (1/b)M_aO_b \qquad (III)$$

where both the metal M and its oxide M_aO_b are either solid or liquid.

The standard Gibbs free energy of formation, $\Delta G^{\circ}_{M_aO_b}$, for the metal oxide, is negative for a stable oxide; the greater magnitude indicates greater stability. Values of $\Delta G^{\circ}_{M_aO_b}$ have been extensively tabulated or plotted against temperature (Ellingham diagram) in the literature. An Ellingham diagram assumes that the heat capacities of the products and reactants are constant even though temperature may change. Thus, a plot of the standard free energy of formation versus temperature results in a straight line. The p0 value corresponding to the O_2 partial pressure that would result from the above reaction is:

$$p0^{\circ}_{M_aO_b} = \frac{(\Delta G^{\circ}_{M_aO_b}/b)}{4.575\ T} \qquad (8)$$

where the Gibbs free energy is in cal/mol.

The oxide M_aO_b will tend to be reduced if

$$p0 > p0^{\circ}_{M_aO_b} \quad \text{or} \quad e0 < \Delta G^{\circ}_{M_aO_b}/b \qquad (9)\ (10)$$

and conversely, the metal will be oxidized if

$$p0 < p0^{\circ}_{M_aO_b} \quad \text{or} \quad e0 > \Delta G^{\circ}_{M_aO_b}/b \qquad (11)\ (12)$$

Figure 4 shows the variation in p0 value with temperature for two different types of gases of varying oxygen contents.

The dash lines of this figure represent the values of $p0^{\circ}_{Cr_2O_3}$ and $p0^{\circ}_{NiO}$ as obtained from Ellingham diagrams. At 1000°C, a hydrogen atmosphere containing 100 atomic ppm of oxygen will barely reduce Cr_2O_3; this oxygen content corresponds to a dew point of -45°F. On the other hand, NiO can be reduced by almost any hydrogen-containing atmosphere short of pure steam, and could be reduced above 1300°C by an inert atmosphere containing 1 atomic ppm oxygen. This example can be extended to develop criteria which describe the oxidation or reduction of alloys or mixed oxides, as well as the solubility of oxygen in solid

or liquid metals and alloys.

Measurement Accuracy, Gain and Errors

The output of an oxygen sensor is usually in error by a relatively constant voltage bias, except where a high net ion current flows through the cell electrolyte. In the latter case, the composition of the oxygen lean side may change appreciably. Assuming that this condition is avoided, the output voltage error is due to uncertainties in the measurement of the reference O_2 pressure, uncertainties in the measurement of the cell temperature, and temperature non-uniformities in the cell. The combined effect of these sources of error is of the order of 10 mv, if nothing is done to optimize accuracy.[5] Since the oxygen sensor has a response of about 56 mv per decade of O_2 partial pressure at 850°C, this voltage error will lead to an absolute measurement accuracy of about 0.2 decade in O_2 partial pressure. However, it is possible to reduce the effect of the cell voltage error to about 1 mv.

Where the logarithm of the O_2 partial pressure is measured below 10^{-15} atm., an absolute accuracy of 0.2 decade leads to at least a relative percentage accuracy of

$$100 \left[\log (2) - \log (1)\right] / 15 \text{ or } 3\%$$

There is not much point in trying to obtain a better relative accuracy in this case. But measurement of the oxygen content in an O_2/diluent gas mixture, leads to a relative percentage accuracy of

$$100 (2z - z) / z \text{ or } 100\%$$

Obviously, some method of eliminating or compensating for the voltage error of the sensor must be used to obtain accurate measurements at higher oxygen pressures. A calibration technique defined by Spacil[5] compensates for the error.

From Equation (1), the term CRT/4F may be expressed as a linear function of temperature:

$$\frac{CRT}{4F} = 55.695 + 0.049588 \ (^{o}C\text{-}850)$$

This corrects for any deviation between the set point and actual temperatures. The latter may be read from a thermocouple attached directly to the electrolyte.

The function of voltage versus oxygen concentration is a straight line on a semi-log plot. The computed plot can be replaced by a similar plot of identical slope which passes through

the point measured with the calibrating gas. Pure oxygen is an especially convenient calibrating gas. A measurement system equipped with an adjustable zero and adjustable range should be used.

The following equation may be derived to define the relative error:

$$\frac{Z^* - Z}{Z^*} = \frac{\Delta P}{Pc} + C\,\frac{\Delta T}{T^*}\,\log\left(\frac{Z^*}{Zc}\right) \qquad (13)$$

where

Z = actual oxygen content

Z^* = indicated oxygen content

Zc = oxygen content of calibration gas

T^* = measured cell temperature

$\Delta T = T - T^*$

T = actual cell temperature

Pc = total pressure of calibrating gas

ΔP = difference in total pressure between sample and calibration gases

$\Delta P/Pc$ and $\Delta T/T^*$ are small.

Equation (13) applies to any calibrating gas.

The pressure contribution to error remains constant, while the temperature contribution increases as the measurement is extended to lower O_2 contents.

With the barometric pressure varying 3% from the mean, the first term is $\pm$ 0.03. With an assumed temperature measurement error of 2°C at 850°C, the value of $C\,\Delta T/T^*$ is $\pm$ 0.0039.

The following table shows the maximum relative error in Z^*, as a percentage:

Z^*	Max. % Error
100%	3.0
1%	3.8
100 ppm	4.6
1 ppm	5.4

Analyzed gas mixtures were used to check the performance of the system after a direct calibration procedure. Typical results were as follows:

Analysis % O_2	Measured % O_2
4.93	4.9
0.515	0.51

The cell gain may be defined as the relative change in voltage affected by a relative change in some other parameter. Take the ratio of the oxygen content of the sample gas to that of the calibrating gas:

$$E_o = \frac{CRT}{4F} \log \frac{[O_2]}{[O_2]_R} = \frac{RT}{4F} \ln \frac{[O_2]}{[O_2]_R}$$

let $X = [O_2]/[O_2]_R$ and $C' = \frac{RT}{4F}$

$$E_o = C' \ln X$$

$$\frac{dE_o}{dX} = \frac{C'}{X} \quad \text{and} \quad \frac{dE_o/E_o}{dX/X} = \frac{1}{\ln X}$$

when $X \longrightarrow 1$, $\frac{1}{\ln X} \longrightarrow \infty$

Thus, maximum gain is obtained when the sample gas and the reference gas are of the same O_2 partial pressure.

Industrial Applications

What aspects of the composition are important and how will they be used? There are situations where the composition itself is important, in order to obtain a measurement of the change in the amount of a particular component. There are situations where the gas composition is used to measure a more abstract property of the gas, generally by computing. Either situation may require fewer than the minimum number of independent measurements required to completely establish the composition. Table I summarizes a number of metallurgical processes in which oxygen is involved. When the O_2 content is above 0.1%, the other species present act as diluents. Paramagnetic oxygen sensors and combustion meters have been commonly used to control processes with this level of O_2. Only recently has the EMF sensor concept been used for this purpose on a combustion process [6]. Combustion control on direct fired soaking pits is a good example of an application for the EMF sensor. The optimum control of combustion would be to regulate the excess air to about 5 $\pm$3% (or excess O_2 to 1 $\pm$.5%) in the combustion gases. Since soaking pits

usually use a fuel of varying composition (coke oven gas, blast furnace hot gas, and natural gas), a feedback measurement of the stack gas 0_2 content has been used to control the burner air/fuel ratio. The low 0_2 content and the particle laden nature of the gas sample makes the application of paramagnetic sensors only marginal. On the other hand, EMF sensor measurements on such a gas have the computed characteristics shown in Figure 5. During tests on a bottom fired soaking pit (Figure 6), we found the logarithmic response of the EMF sensor, as well as its insensitivity to particles in the gas phase, extremely desirably characteristics. The EMF sensor has also been used successfully on the BOF process in series with other instruments to allow calculation of the way lance oxygen is being used.

The EMF sensor is very well suited to determination of the stoichiometric combustion point, SCP. The drastic change in the equilibrium 0_2 partial pressure of combustion products as the air/fuel ratio is varied through the SCP leads to a step-like function in the output of the EMF sensor. (Figure 7). The SCP is an example of how one particular aspect of composition in multicomponent gases can be determined by an EMF sensor without additional information needed.

When the free 0_2 is less than 10^{-6} atmospheres, the other species in the gas usually determine, by reaction, the equilibrium 0_2 partial pressure. Measurement of other species of the gas phase (such as CO, CO_2, H_20 and H_2) have been and are being used for control purposes. Combinations of these measured or assumed species concentrations are being used to show effect on the product (oxidized, reduced and carburized). Equilibrium is assumed so that reactions I and II hold.

Dewpoint, infrared and paramagnetic instruments do not cause the gas to attain equilibrium at a specifically known temperature as does the EMF sensor. Therefore, the use of the former types of sensors are all based on an assumption that does not apply to the EMF sensor. In many cases, measurements made by these are being used to indirectly determine the 0_2 partial pressure of the gas by use of relationships such as Reaction I and II. For these applications, the EMF sensor should be used since it measures this quantity directly under known temperature and pressure conditions.

Annealing of metal product under protective atmospheres is a prime example of the use of indirect measurements to determine the 0_2 partial pressure of an atmosphere. Prior to development of the EMF sensor, instruments were not available to measure down to 10^{-27} atmospheres oxygen (equivalent dewpoint of $-80^{o}F$) as required for applications such as the bright annealing of stainless steel strip.

Industrial tests have been carried out on a continuous annealing line for tinplate. An EMF sensor monitored the oxygen partial pressure along the length of the furnace (Figure 8). These tests were used to diagnose those conditions (air inleakage, too high an 0_2 partial pressure in the cooling zone, etc.) that cause strip oxidation. The generation and control of atmospheres for obtaining a specific carbon potential and/or for carburizing appears to be another area of application for the EMF sensor. The range of 0_2 partial pressures encountered, from 10^{-12} to 10^{-20} atmospheres, would lead to a voltage of from 700 to 1200 mv for a cell operating at 850°C with an air reference electrode. Digital simulations of the gas generator, drier, scrubber, and annealing furnace (Figure 9) indicate that the EMF sensor measurement will be sensitive to changes in air/fuel ratio, dryer and scrubber efficiencies, and carburizing load (Figure 10).

Since the atmosphere in the furnace will come to equilibrium, EMF sensor measurements should allow correlation with the carbon potential of the gas in the metal treating furnace. We have had an oxygen sensor monitoring a hydrogen treating furnace at the Instrument Department for over four months. A sequence sampling system analyses the supply hydrogen, inlet mix and furnace effluent in turn. An argon flush is injected between samples. For the first time, we are getting to understand what is really going on in the process. Not surprisingly, the first important information defined gross leaks in the system. These were in the piping, valves, and in the furnace itself. Malfunctions of the dryer system on the hydrogen supply were also revealed. We are now able to evaluate the parameters of the process itself.

Signal Conditioning

An EMF sensor measures equilibrium 0_2 partial pressure, but some of its readout scale options are linear in p0 and e0 values. Also, an EMF sensor measures the 0_2 partial pressure at its cell temperature. The p0 and e0 criteria for oxide reduction or formation, however, must be applied at the processing temperature. Appropriate expressions for obtaining correct p0 and e0 values at the processing temperature are as follows:

$$p0_s = p0 - \gamma \qquad (13)$$

$$e0_s = \alpha\, e0 + \beta \qquad (14)$$

where p0 and e0 with no subscript are the values measured by the EMF sensor and $p0_s$ and $e0_s$ are the corrected values for the processing temperature. The $\alpha = T_s/T$ where T_s and T are the processing and sensor temperatures in °K. The adders γ and β depend on the nature of the gas sample. For 0_2/diluent atmo-

spheres, both are zero. For CHOD* atmospheres, these quantities depend on the exact composition of the sample but average values can be used with no more than about a 5% error in corrected p0 and e0 value. The average values depend only on sensor and processing temperatures. (Figures 11 and 12). The EMF sensor has a unique ability, therefore, to measure p0 and e0 values in a wide variety of atmospheres , with only a very general knowledge of the atmosphere composition required in order to apply any necessary temperature correction. This correction can be wired into the sensor.

Summary and Conclusion

The theory of solid electrolyte oxygen sensors has been presented. The operating principle is based on the Nernst Equation. Various references may be used; air and nickel-nickel oxide are discussed. Hardware is described. Diluent systems involving nitrogen, for example, and reactive systems involving hydrogen and/or carbon monoxide are discussed. Logarithmic functions such as p0 and e0 are useful for defining oxygen contents in these systems since they provide wide range coverage in a simple manner.

Industrially, the solid electrolyte (zirconia) oxygen sensor may be applied to combustion, soaking pits, BOF, direct reduction, reheat furnaces, annealing furnaces, and carburizing furnaces. There are other applications but are not included in this paper.

Proper signal conditioning is required to accommodate the logarithmic voltage characteristics of the zirconia cell. Bias for differences between sensor and process temperature can be readily provided.

The normal uncorrected cell voltage error is about 10 mv. It is possible to reduce the effect of this error to about 1 mv by proper calibration and using measured instead of set point cell temperature. The calculated maximum percent error (of the value) is 3% at 100% O_2, 3.8% at 1% O_2, and 5.4% at 1 ppm O_2. Measured errors were 0.6% at 4.9% O_2, and 1% at 0.5% O_2. For reactive systems, where the oxygen partial pressure is less than 10^{-15} atm., a 3% (of the value) accuracy is indicated.

Cell gain is defined as the ratio of the relative change in voltage to the ratio of the oxygen content of the sample gas to that of the calibrating gas. A mathematical analysis derives an equation which indicates that maximum cell gain is

---

* CHOD - an atmosphere containing carbon, hydrogen, oxygen and inert diluent.

obtained when the sample gas and the reference gas are of the same oxygen partial pressure.

In conclusion, the solid electrolyte oxygen sensor provides a superior, accurate, reliable, and repeatable measurement for practical applications in industrial and research areas.

Bibliography

1. H.S. Spacil, The Solid Electrolyte Oxygen Sensor, I - Principles and Applications.

2. H. Schmalzried, Zeit, für Elektrochemie 66, 572 (1962).

3. H.S. Spacil and D.L. Schroeder, The Importance of Oxygen Partial Pressure Measurements in the Metals Industry, A.I.M.E. Fall Mtg., Cleveland, Oct. 1967.

4. D.W. White, Progress in High-Temperature, Zirconia-Electrolyte Cell Technology at General Electric. Electrochemical Society Mtg. Fall 1968, Montreal.

5. H.S. Spacil, The Solid-Electrolyte Oxygen Sensor II High Precision Measurements in O_2/Diluent Gases, GE R&DC Report No. 67-C-450, Dec. 1967.

6. W.M. Hickam and J.F. Zamaria, Inst. and Control Syst. 40, 87 (Aug. 1967).

TABLE I

PROCESS	APPLICATION	INSTRUMENT NOW USED
BOF	Stack Gas Analysis (pO, CO/CO_2 Ratio)	Infrared CO & CO_2 Paramagnetic O_2
Blast Furnace	Stack Gas Analysis (CO/CO_2, H_2/H_2O Ratios)	Infrared CO & CO_2 Hygrometers
Direct Reduction	Reducing Gas Composition (eO, CO/CO_2, H_2/H_2O Ratios)	Infrared CO & CO_2 Hygrometers
Reheat Furnaces	Combustion Control (pO, Excess O_2)	Paramagnetic O_2 Combustion Meters
Annealing Furnaces	Atmosphere Control (eO, pO, CO/CO_2, H_2/H_2O ratios)	Infrared CO & CO_2 Hygrometers
Carburizing Furnaces	Carbon Potential Control (pO, CO/CO_2, Ratio)	Infrared CO & CO_2 Hygrometers

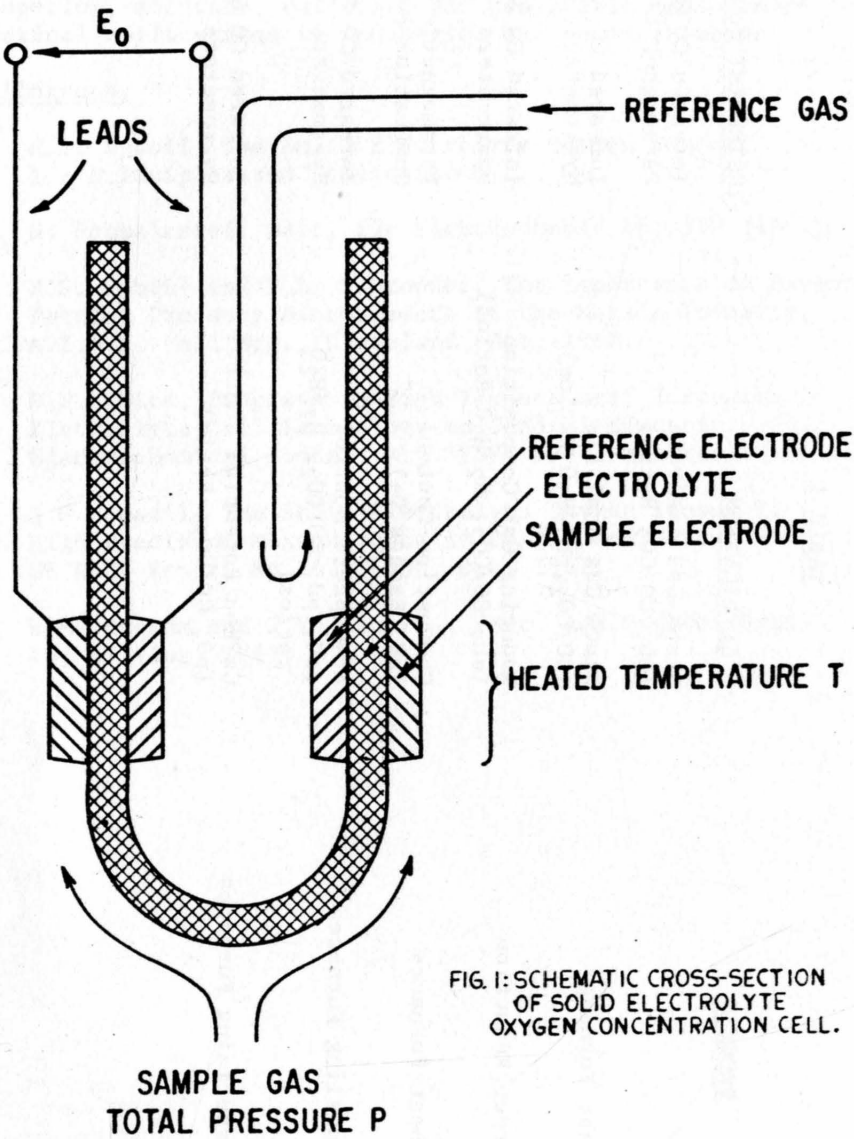
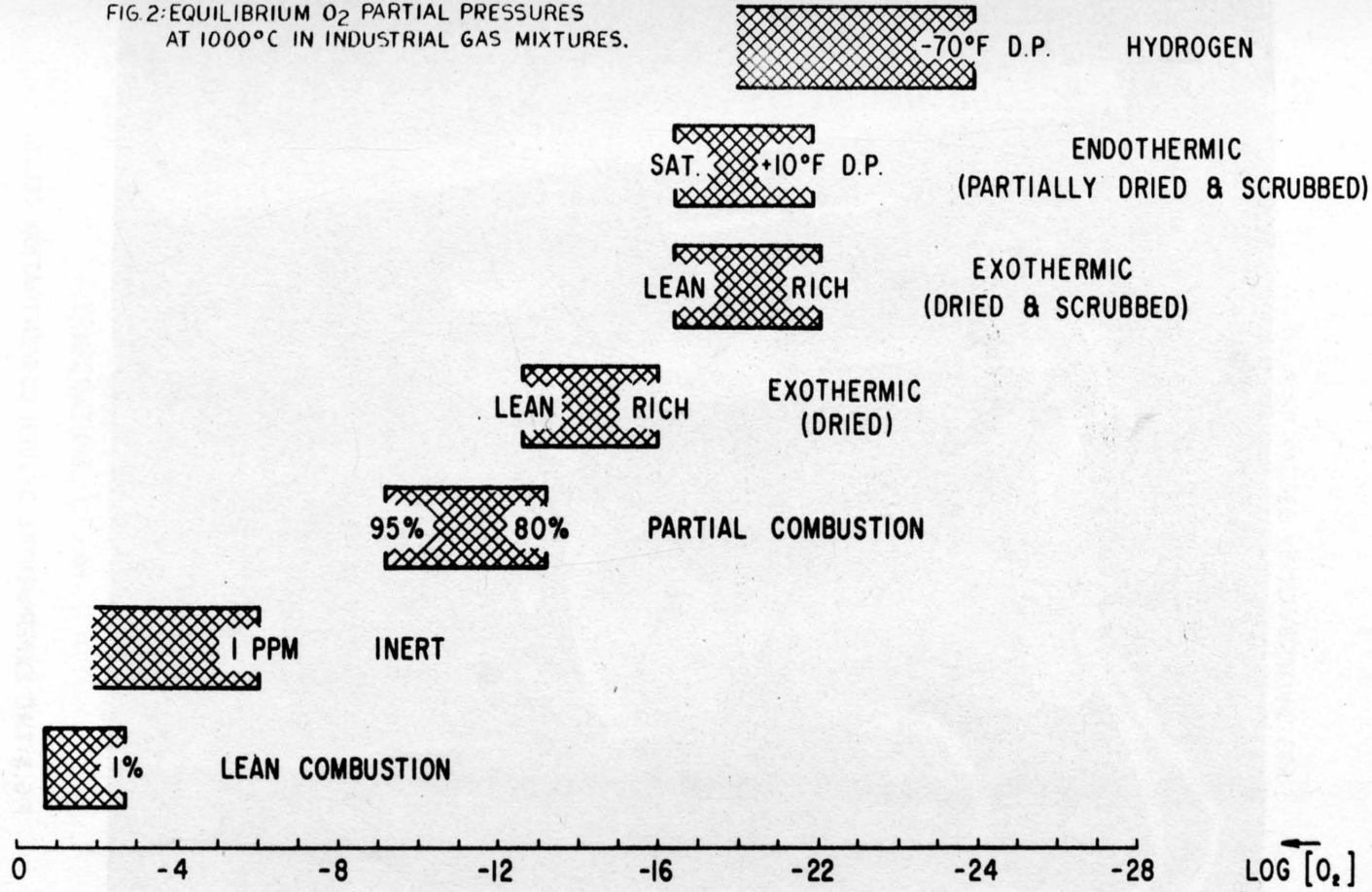


FIG. 1: SCHEMATIC CROSS-SECTION OF SOLID ELECTROLYTE OXYGEN CONCENTRATION CELL.

FIG. 2: EQUILIBRIUM O_2 PARTIAL PRESSURES AT 1000°C IN INDUSTRIAL GAS MIXTURES.



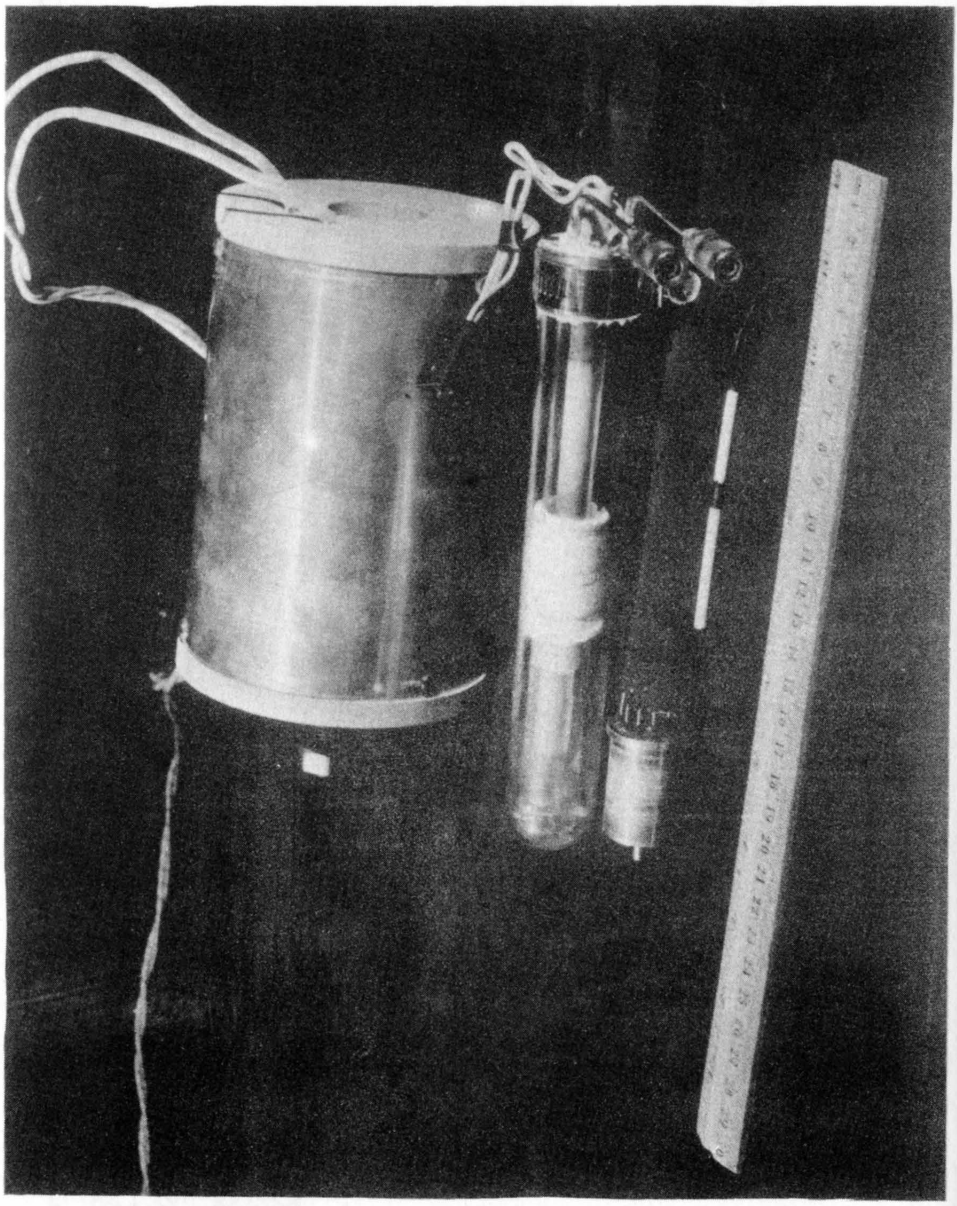


FIG.3: TWO EXPERIMENTAL OXYGEN CONCENTRATION CELLS.

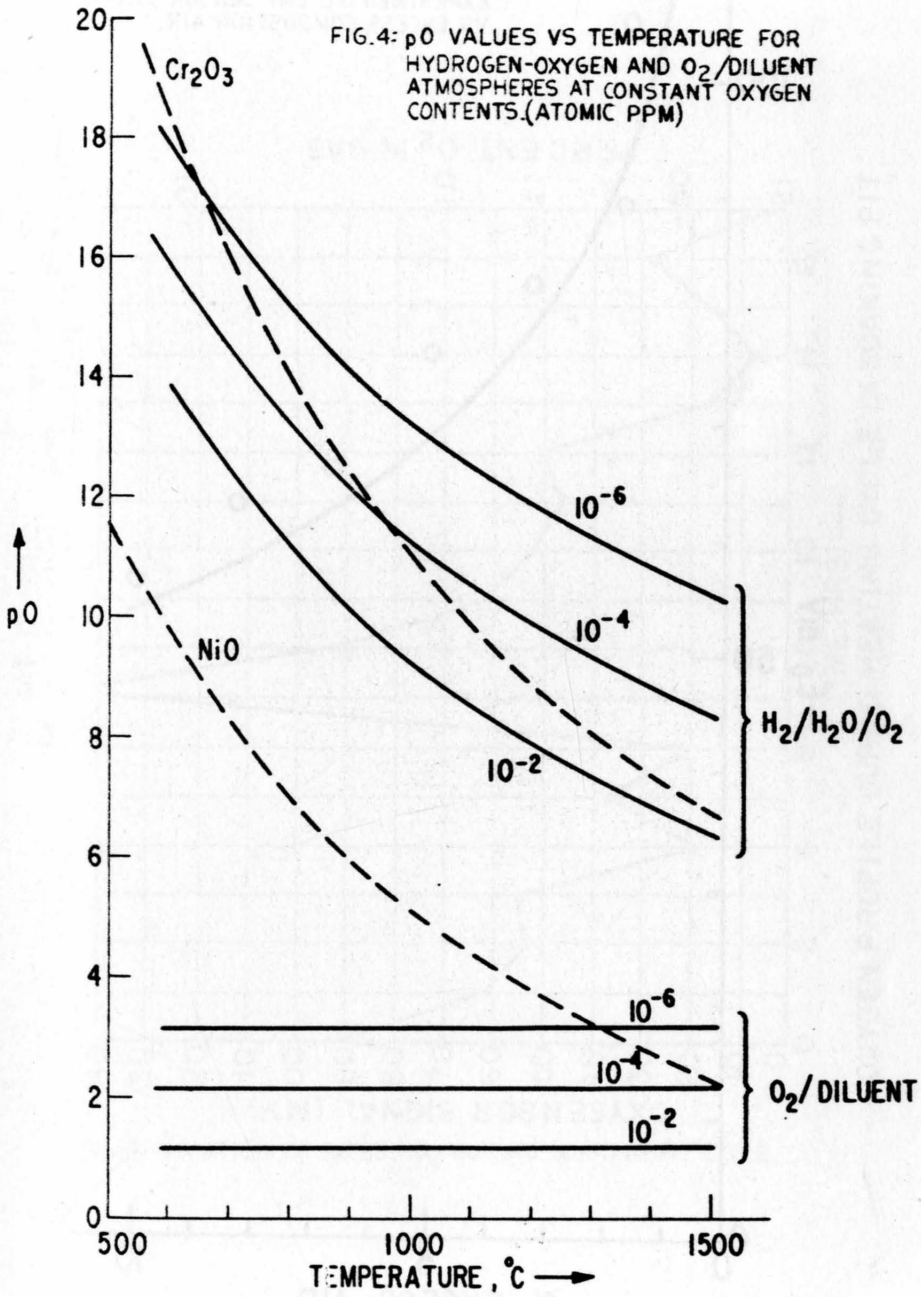
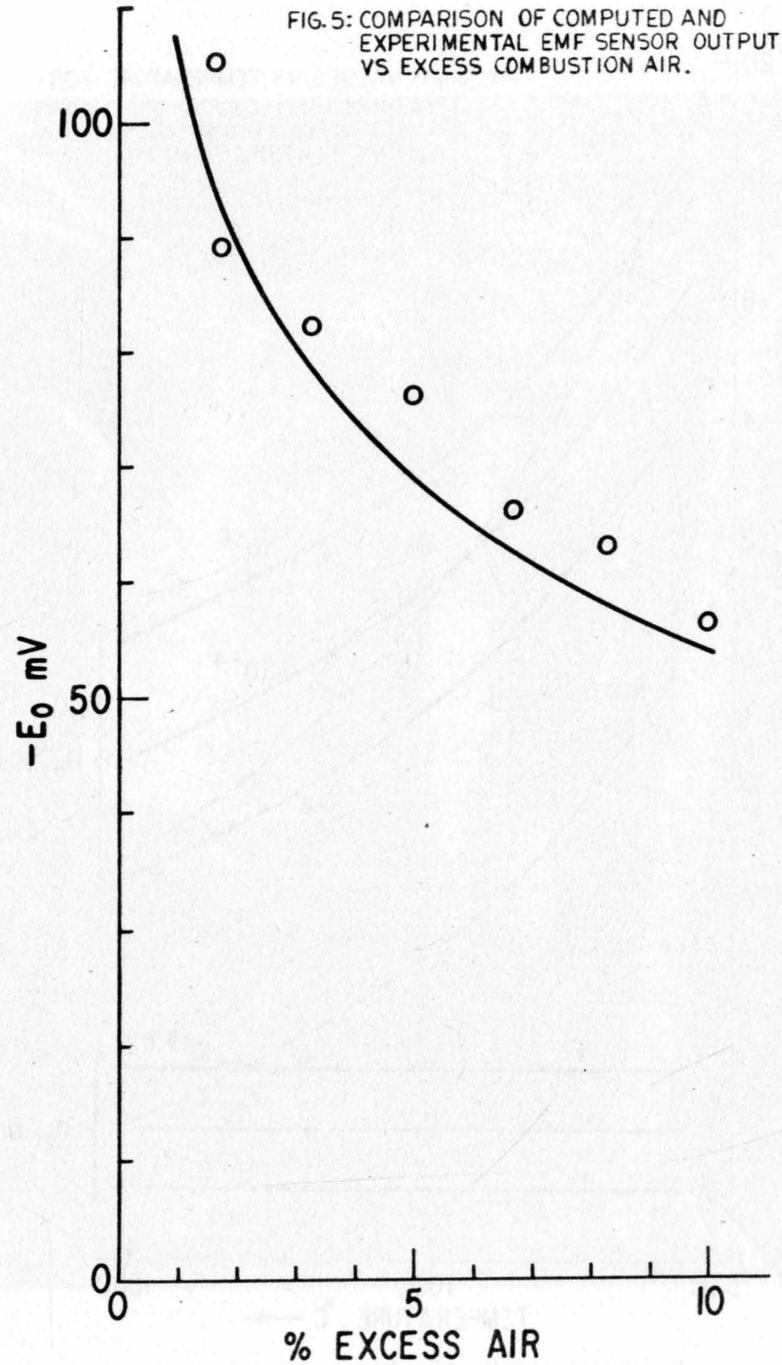


FIG.4: pO VALUES VS TEMPERATURE FOR HYDROGEN-OXYGEN AND O_2/DILUENT ATMOSPHERES AT CONSTANT OXYGEN CONTENTS.(ATOMIC PPM)
Cr_2O_3
NiO
10^{-6}
10^{-4}
10^{-2}
$H_2/H_2O/O_2$
10^{-6}
10^{-4}
10^{-2}
O_2/DILUENT
pO
TEMPERATURE, °C
20
18
16
14
12
10
8
6
4
2
0
500
1000
1500

FIG. 5: COMPARISON OF COMPUTED AND EXPERIMENTAL EMF SENSOR OUTPUT VS EXCESS COMBUSTION AIR.



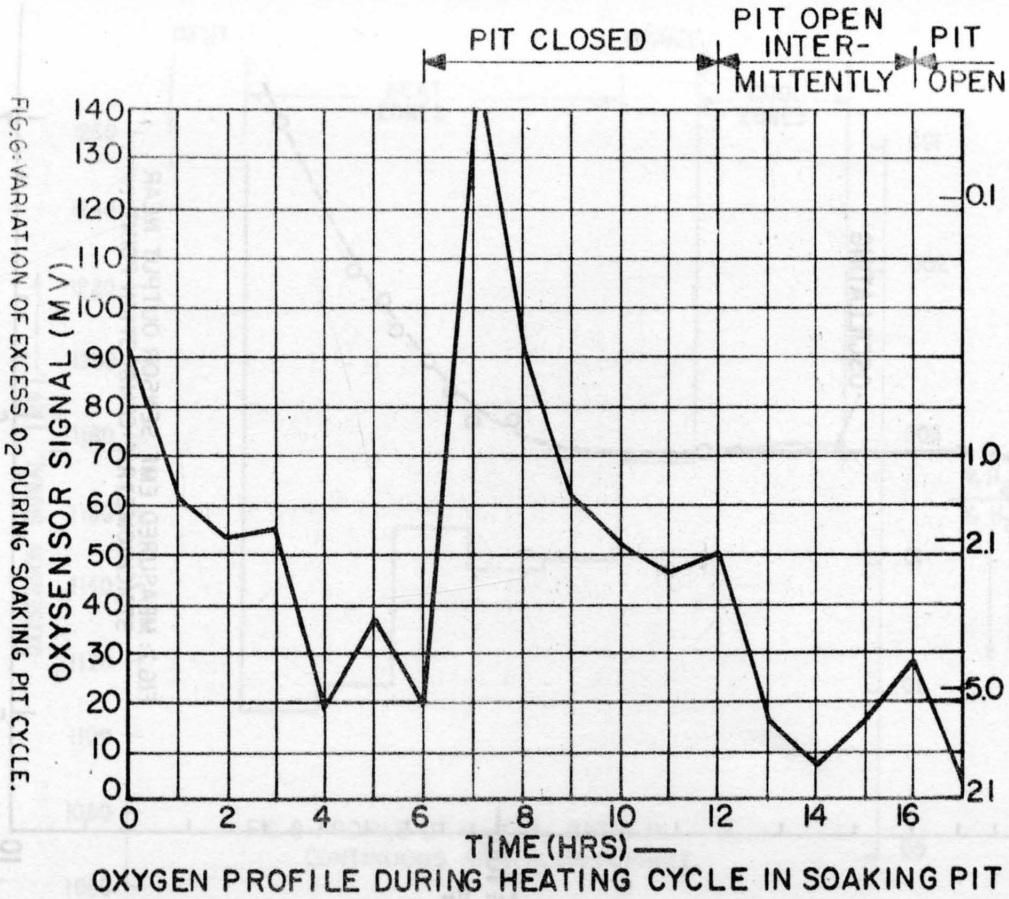


FIG.6: VARIATION OF EXCESS O_2 DURING SOAKING PIT CYCLE.

OXYGEN PROFILE DURING HEATING CYCLE IN SOAKING PIT

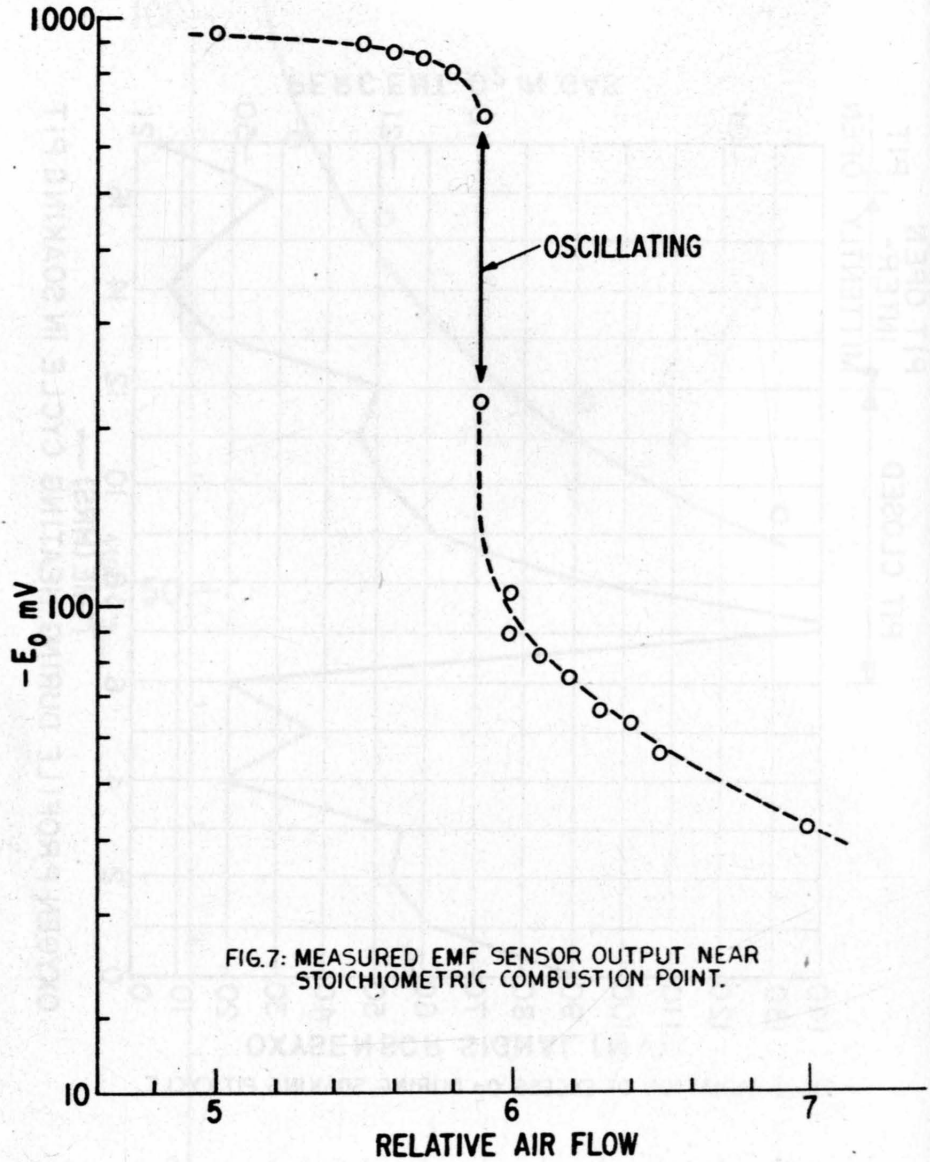


FIG.7: MEASURED EMF SENSOR OUTPUT NEAR STOICHIOMETRIC COMBUSTION POINT.

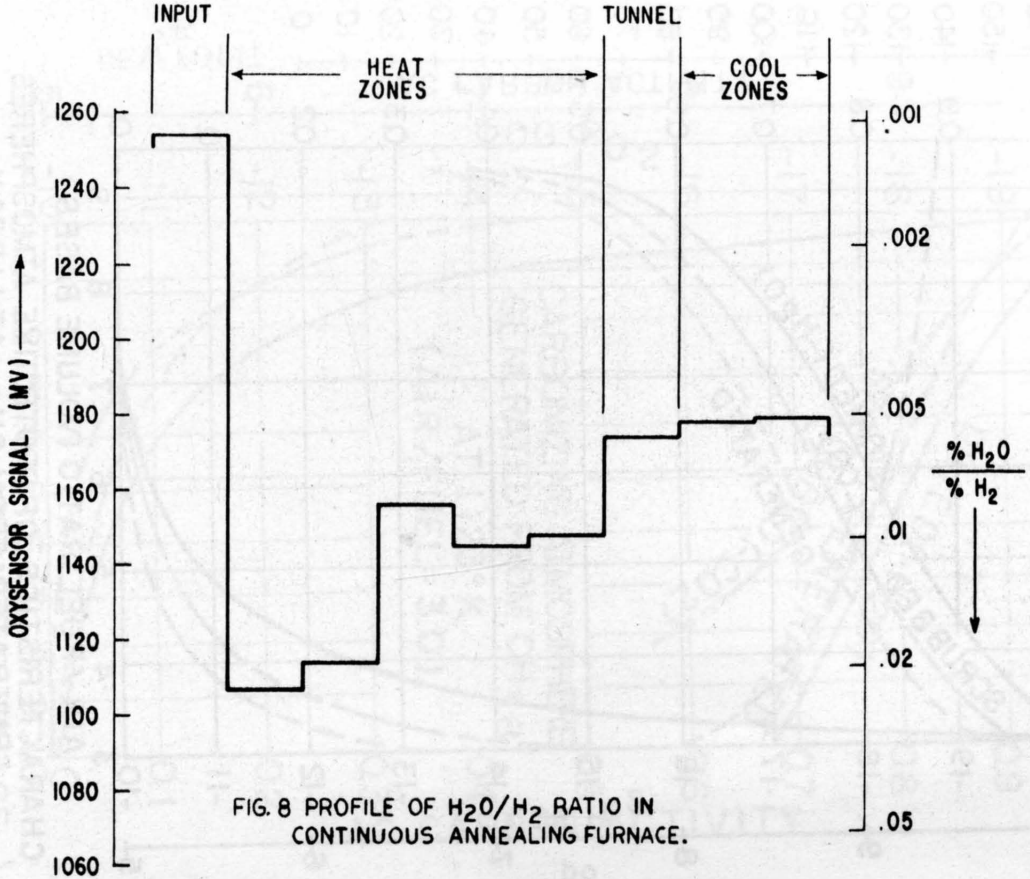
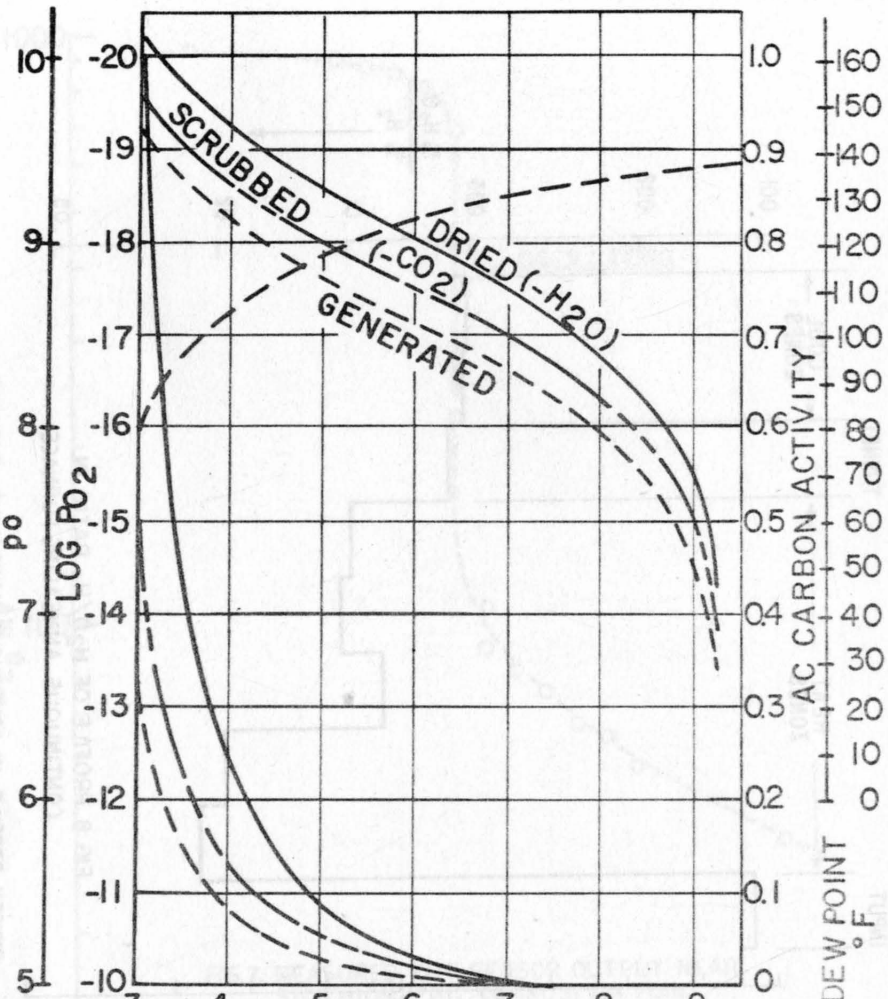


FIG. 8 PROFILE OF H_2O/H_2 RATIO IN CONTINUOUS ANNEALING FURNACE.

OXYGEN PROFILE IN CONTINUOUS ANNEALING FURNACE



CHARACTERISTICS OF PROTECTIVE ATMOSPHERES GENERATED FROM $CH_{3.8}$ AT 1123° K

FIG. 9: COMPUTED PROPERTIES OF ATMOSPHERE GENERATOR OUTPUT VS AIR/FUEL RATIO.

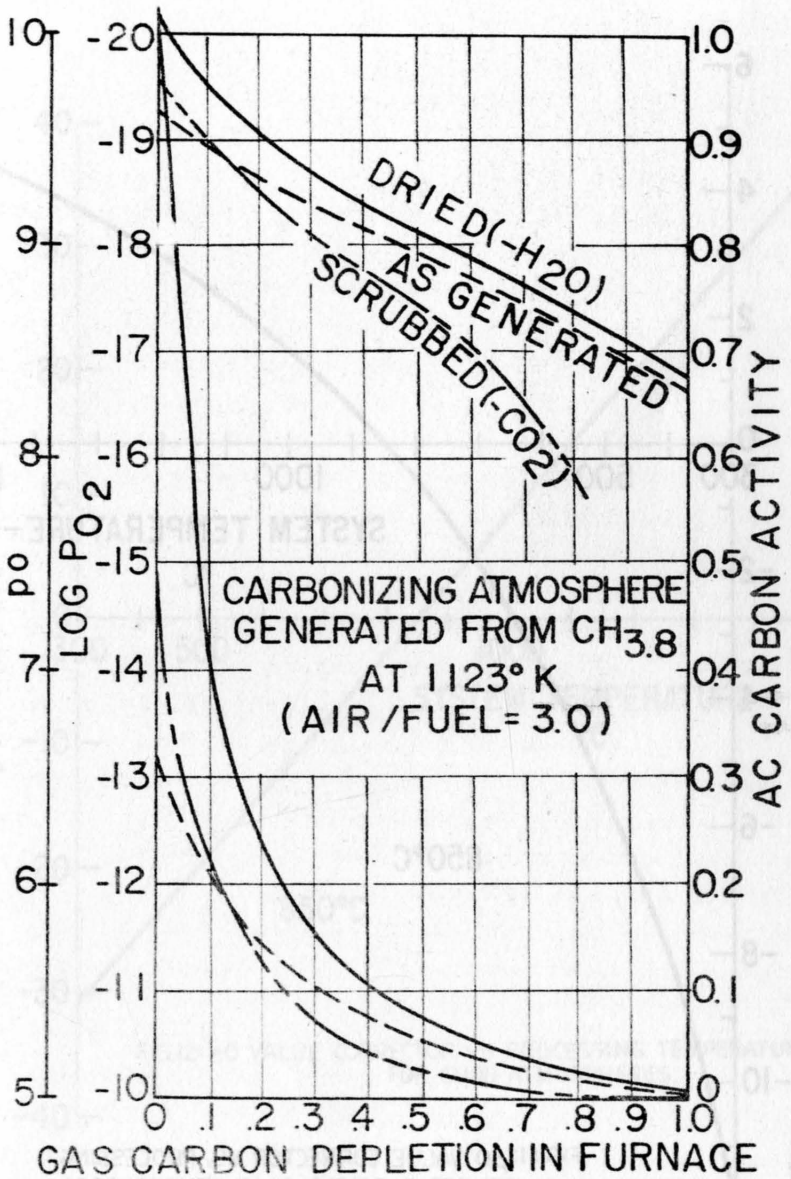


FIG.10: COMPUTED PROPERTIES OF ATMOSPHERE GENERATOR OUTPUT VS CARBON DEPLETION.

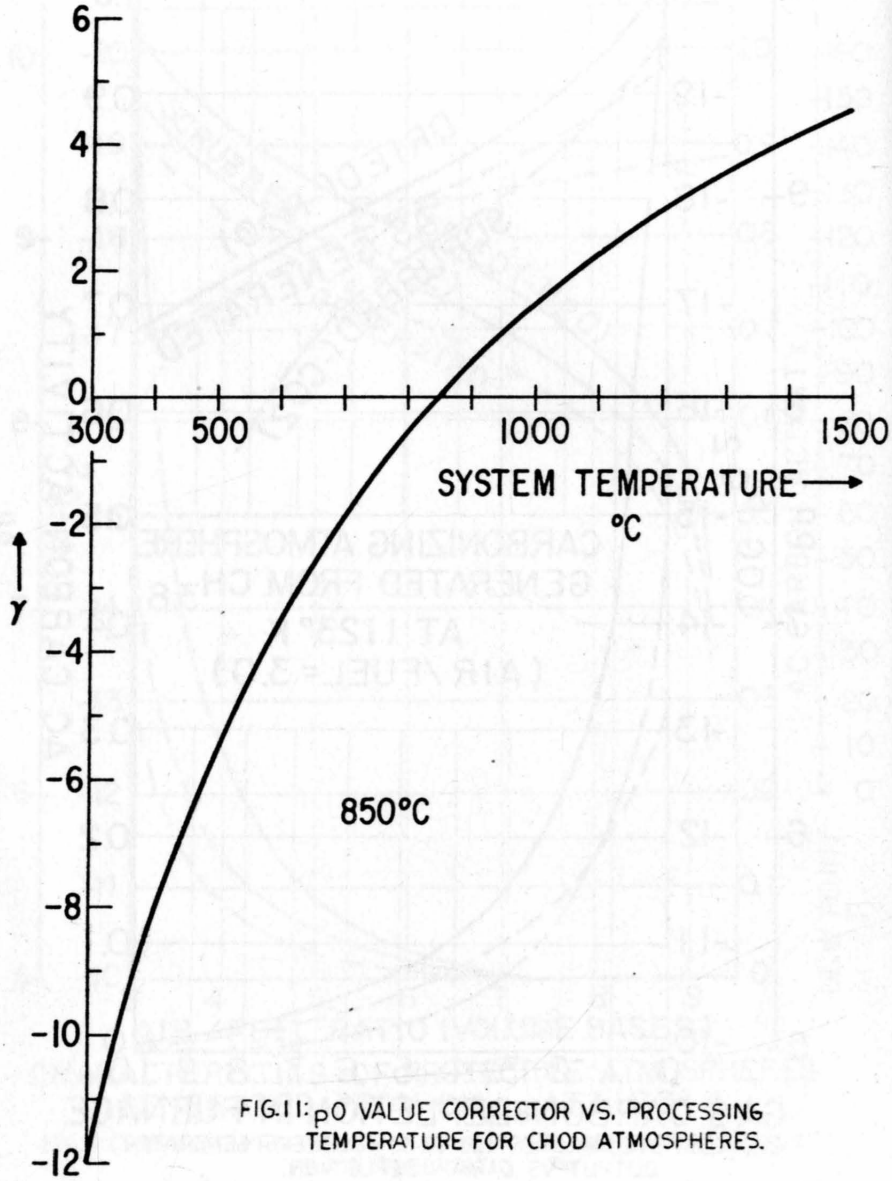


FIG.11: pO VALUE CORRECTOR VS. PROCESSING TEMPERATURE FOR CHOD ATMOSPHERES.

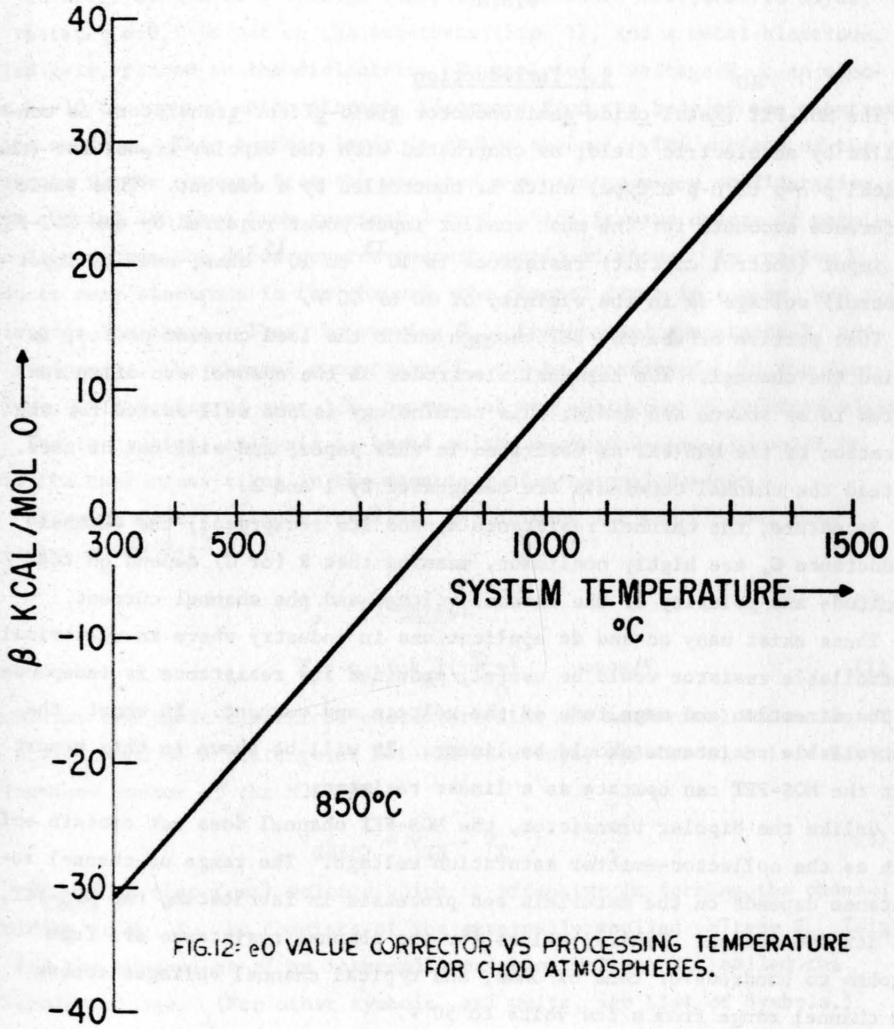


FIG.12: eO VALUE CORRECTOR VS PROCESSING TEMPERATURE FOR CHOD ATMOSPHERES.

VOLTAGE-CONTROLLED MOS-FET RESISTOR

H. F. Storm
General Electric Research & Development Center
Schenectady, N. Y.
U.S.A.

1. Introduction

The MOS-FET (metal-oxide-semiconductor field-effect-transistor) is controlled by an electric field, as contrasted with the bipolar transistor (the typical p-n-p or n-p-n type) which is controlled by a current. This basic difference accounts for the much smaller input power required by the MOS-FET. Its input (control circuit) resistance is 10^{13} to 10^{15} ohms, and its input (control) voltage is in the vicinity of 10 to 40 V.

That portion of the MOS-FET through which the load current passes, is called the channel. The terminal electrodes of the channel are often referred to as source and drain; this terminology is not well suited for the operation of the MOS-FET as described in this paper, and will not be used. Instead the channel terminals are designated by 1 and 2.

By nature, the channel resistance R, and its reciprocal, the channel conductance G, are highly nonlinear, meaning that R (or G) depend on the magnitude and polarity of the channel voltage and the channel current.

There exist many ac and dc applications in industry where an electrically controllable resistor would be useful, provided its resistance is independent of the direction and magnitude of the voltage and current. In short, the controllable resistance should be linear. It will be shown in this report that the MOS-FET can operate as a linear resistor.

Unlike the bipolar transistor, the MOS-FET channel does not contain emfs such as the collector-emitter saturation voltage. The range of channel resistance depends on the materials and processes in fabricating the MOS-FET, and its dimensions. Typical values for the channel resistance are from Megohms to hundreds or tens of ohms, and typical channel voltages across the channel range from a few volts to 50 V.

These and other MOS-FET properties suggest it for many applications, for instance, for long-time timers, variable R-C time-constant circuits, voltage controlled attenuators, adaptive controls, multipliers, flip-flops, memories, and for amplification and modulation.[1-9]

2. Causes of Nonlinearity of MOS-FET Channel Resistance

The MOS-FET consists of a substrate, also called body, and 2 diffusions

of the opposite carrier type (Fig. 1). If terminals 1 and 2 are connected to an ac supply voltage, a small alternating current flows which is controlled essentially by the reverse characteristics of the 2 diodes connected back-to-back.

In order to obtain a voltage controlled resistor, a dielectric layer, for instance SiO_2, is put on the substrate (Fig. 2), and a metal electrode, called gate, placed on the dielectric. By applying a voltage V_{GB}, an electric field is created which attracts electrons from the body of the substrate to its surface. This surface layer is called channel. The surface of the substrate is now changed from the previous n-p-n to an n-n-n configuration, and a current can pass from terminal 1 to 2, which is many orders of magnitude larger than the diode reverse current mentioned above. Increasing V_{GB} produces more electrons in the channel, the channel depth increases, and its resistance decreases. Thus, by varying V_{GB}, the channel resistance R, and its reciprocal, the channel conductance G, can be controlled. In the derivations, G is preferred over 1/R, because of the avoidance of fraction signs.

The theoretical analysis is based on the n-channel type, in order to avoid the many minus signs in the equations of p-channel devices.

Assuming several simplifications, one obtains for the channel conductivity G_o [10,11,12]

$$G_o = F|V_{GB(E)}| \tag{1}$$

$$F = \varepsilon_o (\bar{\mu} k_O)(\frac{w}{l \cdot t_O}) \qquad \text{mhos/V} \tag{2}$$

F contains the basic electrical characteristics and mechanical dimension of the active part of a rectangular MOS-FET structure, and hence, F is a basic performance factor of the MOS-FET.

$$V_{GB(E)} = V_{GB} - V_T \tag{3}$$

$V_{GB(E)}$ is that (net) voltage which is effective in forming the channel according to Eq. 1. It consists of the externally applied voltage V_{GB} (Fig. 2), and the equivalent of an internally existing voltage V_T, called the threshold voltage. (For other symbols, and units, see List of Symbols.)

Next, we make provision for current $I_{1,2}$ by establishing a voltage $V_{1,2}$ between terminals 1 and 2 (Fig. 3). The terminals 2 and B are directly connected in order to define the voltage between substrate and channel. The question is the magnitude of $I_{1,2}$.

With a voltage $V_{1,2}$ between terminals 1 and 2, the potentials along the channel will assume intermediate values. These potentials will modify the electric field which was previously established by $V_{GB(E)}$, and hence, the conductance of the channel will differ from G_o of Eq. 1. This modifying

effect is called internal feedback or channel reaction, and can be positive or negative. With the polarities shown in Fig. 3, the channel potential will weaken the electric field set up by $V_{GS(E)}$, and hence, the feedback is negative. This is shown in Fig. 3, where the channel becomes more shallow toward terminal 1. If the polarity of $V_{1,2}$ were opposite to the one shown in Fig. 3, the channel potential would strengthen the electric field set up by $V_{GB(E)}$, the channel would become deeper toward terminal 1, and the internal feedback would be positive.

The next question is to the extent by which the channel conductance is affected by the internal feedback. If a quick guess were allowed, one might say that the feedback voltage is zero at $x = 0$, and $v_{1,2}$ at $x = \ell$, hence, the average effect along the channel is perhaps $V_{1,2}/2$. A more elaborate analysis shows that this answer is exactly right (for the range limited by Eq. 6). In view of the above and Eq. 1, the channel conductance of p- or n-channel devices

$$G = F\left|V_{GB(E)} - \tfrac{1}{2}V_{1,2}\right| \tag{4}$$

with $V_{GB(E)}$ and $V_{1,2}$ having proper signs. In view of Eq. 1, the internal feedback is expressed by the term $-\frac{1}{2}V_{1,2}$, where $V_{1,2}$ is a signed quantity. We note that the channel conductance (Eq. 4) is not constant. Instead, it is a function of the feedback voltage $\frac{1}{2}V_{1,2}$, which, in turn, is derived from the voltage $V_{1,2}$ across the channel. Thus, the channel conductance is nonlinear.

The channel current is

$$I_{1,2} = V_{1,2}\, G = F\, V_{1,2} \cdot \left|V_{GB(E)} - \tfrac{1}{2}V_{1,2}\right| \tag{5}$$

We note that the channel current is a quadratic function of the voltage $V_{1,2}$ across the channel.

If $V_{GB(E)}$ and $V_{1,2}$ have the same sign, the feedback is negative. Then even the maximum feedback voltage $V_{1,2}$ at $x = \ell$ cannot make the channel conductivity negative, and one obtains as boundary for the validity of Eqs. 4, 5, and under conditions of negative, internal feedback:

$$\tfrac{1}{2}\left|V_{1,2}\right| \leq \left|V_{GB(E)}\right| \tag{6}$$

When operating with positive internal feedback, that is, when $V_{GB(E)}$ and $V_{1,2}$ have opposite signs, the condition of Eq. 6 is met inherently, but a new restriction for the validity of Eq. 5 develops. A shunt path to the channel is created from terinal B to 1 through the forward biased p-n junction (Fig. 3). This shunt path will carry a significant current when the above junction is foward biased by more than 0.4 to 0.6 V (for

silicon). It follows that Eq. 5 is valid only if the bypass current is negligible. In addition, the bypass current introduces a nonlinearity to the resistance as seen from terminals 1 and 2.

Further boundaries are established by the effective gate voltage $V_{GB(E)}$:

$$V_{GB(E)} \geq 0 \text{ for p-channel MOS-FET} \tag{7a}$$

$$V_{GB(E)} \leq 0 \text{ for n-channel MOS-FET} \tag{7b}$$

One concludes that the major causes of MOS-FET nonlinearity are

(a) the internal, electrostatic feedback expressed by $\frac{1}{2}V_{1,2}$ in Eq. 4

(b) the bypass current flowing between terminals 1 and B (Fig. 3) when $V_{1,2}$ is negative, or alternating.

3. Linearization of the MOS-FET Channel Resistance

As previously stated, one cause of the nonlinearity is the internal feedback which is expressed by $\frac{1}{2}V_{1,2}$ in Eq. 4. If the MOS-FET is to become linear, the internal feedback must be nullified. This nullification is accomplished by introducing an external feedback voltage having an effect which is opposite and equal to the internal feedback.

With the voltage $V_{1,2}$ and $V_{GB(E)}$ having the same sign (Fig. 4), the internal feedback is $-\frac{1}{2}V_{1,2}$, and hence, a potentiometer P whose resistance midpoint N is connected to the gate G (via the gate control voltage V_G) will provide the desired cancelling effect of $+\frac{1}{2}V_{1,2}$. The net channel conductance G is now a linear function of $V_{G(E)}$

$$G = F\left|V_{G(E)} - \tfrac{1}{2}V_{1,2} + \tfrac{1}{2}V_{1,2}\right| = F\left|V_{G(E)}\right| \quad \text{mhos} \tag{8}$$

Similar to Eq. 3

$$V_{G(E)} = V_G - V_T \quad V \tag{9}$$

where V_G and V_T are signed quantities. Analogous to Eq. 6, there is a boundary for Eq. 8. The argument is that even the maximum internal feedback voltage $V_{1,2}$ cannot make the channel conductivity negative and hence $V_{1,2}$ must be equal to or smaller than the sum of the externally applied gate voltages. Some elementary considerations lead to the boundary for Eq. 8:

$$\left|V_{1,2}\right| \leq 2\left|V_{G(E)}\right| \tag{10}$$

Subject to the limitation expressed by Eq. 10, the channel current $I_{1,2}$ becomes a linear function of the supply voltage $V_{1,2}$, and the effective gate voltage $V_{G(E)}$

$$I_{1,2} = V_{1,2} \cdot F\left|V_{G(E)}\right| \quad A \tag{11}$$

As previously pointed out, the internal negative feedback causes a

reduction of channel conductivity, as graphically indicated by a decreasing channel depth in Fig. 3. The positive external feedback causes a uniform increase of channel depth. The net effect is shown in Fig. 4, where the channel depth still decreases from left to right, but where the total channel depth at any one point is larger than shown in Fig. 3. As a consequence, the channel conductance G of the linearized MOS-FET is now the same as for the case shown in Fig. 2, with $V_{GB} = V_G$:

$$G \equiv G_o \tag{12}$$

As stated before, the current $I_{1,2}$ may significantly and nonlinearly exceed the value shown in Eq. 11 on account of a current component which bypasses the channel by flowing from terminal 1 to B. This bypass current can be stopped by disconnecting terminal B. Then, the substrate will assume a floating potential and thereby somewhat modify the characteristics of the MOS-FET. The results are shown in the following.

First, let us operate the MOS-FET (Fig. 5) with zero external feedback (slider of potentiometer P at K), resulting in the parabola A'-A" of Fig. 6 for $V_G = -12$ where the internal feedback is negative; the section from the origin to A' is less known, and has positive internal feedback.

Then, the slider of the potentiometer is moved to the other extreme position at M. The result is the parabola B'-B". This parabola is centrally symmetrical to A'-A" and thereby indicates symmetry of the MOS-FET.

Now we position the slider of P at the resistance half-point N and thereby introduce external feedback of $\frac{1}{2}v_{1,2}$. The cathode ray display shows a straight line C'-C", as predicted by Eq. 11. One concludes that the channel conductance of the MOS-FET, with external feedback of $\frac{1}{2}v_{1,2}$ and floating substrate, is linear.

Next, we would like to see if the linearized channel conductance G is proportional to the gate voltage V_G (Fig. 5) as predicted by Eqs. 8, 9. For this purpose, several gate voltages ($V_G = -10, -15, -20$ V) are sequentially applied. The results are the practically straight-line voltage-current characteristics shown in Fig. 7. The slopes correspond to the channel conductances G. The 3 conductance values for G are plotted as function of V_G in Fig. 8. The measured points are located on a straight line. From this, we conclude that the proportionality expressed by Eqs. 8, 11 also holds for a MOS-FET with floating substrate. Equation 11 indicates that the linearized MOS-FET may be utilized as a multiplier.

Normally, the substrate is connected to a MOS-FET metallic enclosure. In order to prevent extraneous electric fields from interfering with the operation of the MOS-FET, the enclosure of the MOS-FET package in a grounded

shield is desirable.

4. The MOS-FET as Controllable Resistor in a Phase Shift Circuit

As an example of an ac application and one believed to be novel, a phase shifter will be discussed, where the variable phasing is accomplished by a voltage which controls the channel resistance of a MOS-FET. Phase shift circuits are used on a large scale for the control of thyristors (SCRs) and other devices.

One of the basic phase shift circuits consists of a center-tapped transformer (or a center-tapped, relatively low-ohm resistor) with the secondary voltages $V_{S,1}$, $V_{S,2}$ (Fig. 9). The load is represented by the output resistance R_0, connected between the centertap Z and the common point W of an RC network. By varying the magnitude of resistance R, the phase angle between the output voltage V_0 and the secondary voltage $V_{S,1}$ (or $V_{S,2}$) can be varied.

Replacing the resistor R in Fig. 9 by a linearized MOS-FET, we obtain Fig. 10. The phase angles and the waveshapes have been measured for frequencies from 60 Hz to 200 kHz. The phase angles are shown in Fig. 11 for open load circuit ($R_0=\infty$), and in Fig. 12 for a load resistance for R_0 = 6 k-ohms. The waveshapes of the phase controlled output voltage V_0 is shown in Fig. 13.

5. Conclusions

Within the defined boundaries, the nonlinearity of the MOS-FET channel resistance is essentially due to a) an internal electrical feedback equal to $-\frac{1}{2}V_{1,2}$, and b) a bypass current entering or leaving the MOS-FET through the substrate terminal. The MOS-FET channel resistance becomes linear through the use of an external, electrical feedback which is equal and opposite to the internal feedback, and by disconnecting the substrate terminal. The external feedback is obtained simply by a resistive voltage divider across the channel terminals of the MOS-FET.

As a practical example, the linearized MOS-FET was demonstrated as the variable resistance in an RC phase shift circuit, tested over a frequency range from 60 Hz to 200 kHz, with the upper frequency limit introduced by the test equipment and not by the MOS-FET.

6. List of Symbols

a_i	Cross section of the channel (perpendicular to the direction of current flow from 1 to 2, Fig. 3), cm^2, (Eq. 2)
F	Performance factor of MOS-FET, mhos/V, (Eq. 2)
G	Channel conductance, mhos, (Eq. 4)
G_o	Conductance of channel for $V_{1,2} = 0$, mhos, (Eq. 1)
$I_{1,2}$	Channel current, A dc, (Eq. 5)
k_0	Dielectric constant of gate dielectric ($k_0 \doteq 4$ for SiO_2), number, (Eq. 2)
ℓ	Length of channel, (Fig. 3), cm, (Eq. 2)
t_0	Thickness of gate dielectric over channel, cm, (Eq. 2)
$V_{1,2}$	Voltage across channel, V dc, (Eq. 4)
V_C	Phase-shifter control voltage, (Fig. 10), V
V_G	Gate control voltage (Fig. 4), V
V_{GB}	Voltage between gate and substrate (body), (Fig. 2), V, (Eq. 3)
$V_{GB(E)}$	Effective gate voltage, V, (Eq. 3)
$V_{G(E)}$	Effective gate control voltage, V, (Eq. 9)
V_T	Threshold voltage, V, (Eq. 3)
$v_{1,2}$	Voltage across channel, instantaneous value, V
w	Width of channel, (perpendicular to current flow from 1 to 2), (Fig. 3), cm, (Eq. 2)
ε_o	Permittivity of free space, $8.85\ 10^{-14}$ F/cm, (Eq. 2)
μ	Mobility of majority carriers in the channel, cm^2/Vs, (Eq. 2)

7. Bibliography

1. Martin, T. B., "Circuit applications of the field-effect transistor," Semiconductor Products, Vol. 5, February 1962, p. 33; and March 1962, p. 30.
2. Neu, F. D.., "Voltage controls solid-state nonlinear resistance," Electronics, Feb. 21, 1964, pp. 36-48.
3. Morgan, A. N., "The FET as an electronically variable resistor," Proc. IEEE (letter), June 1966, pp. 892-893.
4. Biloti, A., "Operation of a MOS transistor as a variable resistor," Proc. IEEE (letter), Aug. 1966, pp. 1093-1094.
5. Biloti, A., "A distributed MOS attenuator," Proc. IEEE (letter), April 1967, pp. 562-563.
6. Gosling, W., "Field-effect transistor applications," (book) John Wiley & Sons, Inc., New York, N. Y., 1965.
7. Sevin, L. J., "Field-effect transistors," (book) McGraw-Hill Book Co., New York, N. Y., 1965

8. Eimbinder, J., "FET applications handbook," TAB Books, Blue Ridge Summit, Pa., 1967
9. Storm, H. F., "Field-effect transistor (FET) bibliography," IEEE Trans. on Electron Devices, Vol. ED-14, #10, Oct. 1967, pp. 710-717
10. Hofstein, S. R., and F. P. Heinman, "The silicon insulated-gate field-effect transistor," Proc. IEEE, Vol. 51, September 1963, pp. 1190-1202
11. Johnson, J. E., "Physical processes in insulated-gate field-effect transistors," Solid-State Electronics, Vol. 7, pp. 861-871
12. Crawford, R. H., "MOS-FET in circuit design," (book) McGraw-Hill Book Co., New York, N. Y., 1967

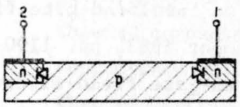


Fig. 1 MOS-FET substrate. Current flow between terminals 1 and 2 is limited by the reverse characteristics of the 2 back-to-back connected diodes.

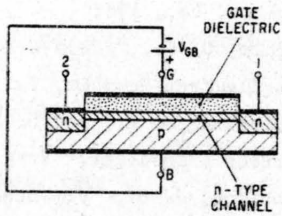


Fig. 2 Basic MOS-FET structure. Dielectric is covered by electrode G, called gate. An inversion layer, called channel, permits current flow between terminals 1 and 2. Terminal B stands for body ≡ substrate. When housed in a metallic can, the terminal B is usually connected to the can.

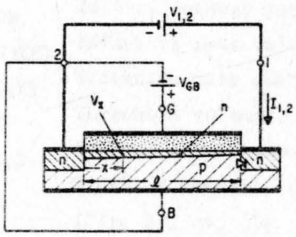


Fig. 3 Usual dc MOS-FET circuit. $V_{1,2}$ causes internal, electrostatic feedback. $V_{1,2}$, as shown, causes negative feedback and channel conductivity is reduced, see narrowing of the channel toward the right. If $V_{1,2}$ is reversed, internal feedback is positive, and the conductivity increases.

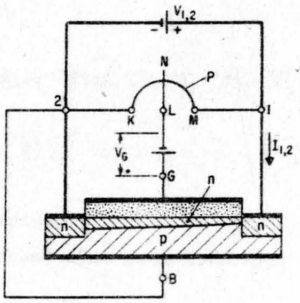


Fig. 4 MOS-FET with external feedback provided by potentiometer P. With negative, internal feedback, channel narrows to the right, but due to external feedback by P, it is wider at all points than channel shown in Fig. 3.

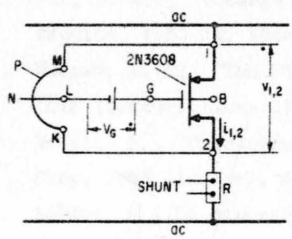


Fig. 5 MOS-FET with external feedback, and floating substrate (terminal B).

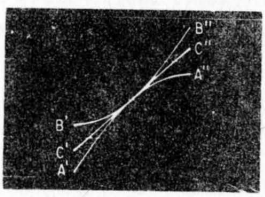


Fig. 6 Current $i_{1,2}$ as function of voltage $v_{1,2}$ of MOS-FET with floating substrate (Fig. 5), for V_G = -12 V, f = 100 Hz.

Horizontal scale, $v_{1,2}$: -5 V/cm
Vertical scale, $i_{1,2}$: -10 mA/cm

A'-A": slider of potentiometer P (Fig. 5) at K
B'-B": " " " " " " M
C'-C": " " " " at midpoint N

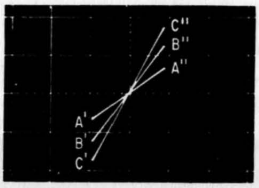


Fig. 7 Current $i_{1,2}$ as function of voltage $v_{1,2}$ of MOS-FET with floating substrate (Fig. 5) and potentiometer at N, f = 100 Hz.

Horizontal scale, $v_{1,2}$: -5 V/cm
Vertical scale, $i_{1,2}$: -10 mA/cm

A'-A'': V_G = -10 V
B'-B'': " = -15 V
C'-C'': " = -20 V

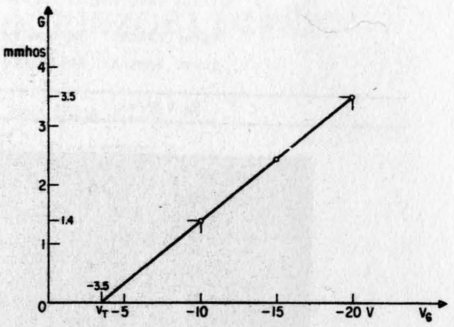


Fig. 8 Dots indicate slopes of the three characteristics of Fig. 7. Slope of straight line of Fig. 8 indicates F = 210 μmhos, intercept indicates V_T = -3.5 V

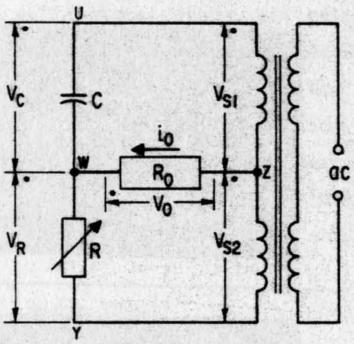


Fig. 9 Basic phase shift circuit. The phase position of the output voltage V_0 is controlled by R.

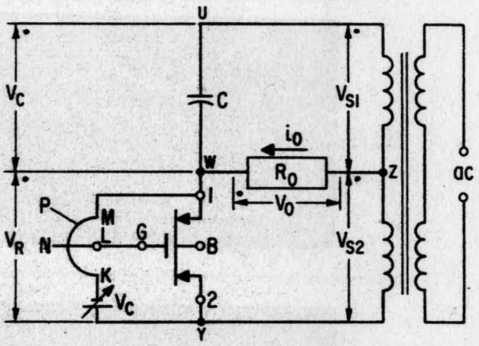


Fig. 10 Phase shift circuit similar to Fig. 9, where a linearized MOS-FET provides the variable resistance R.

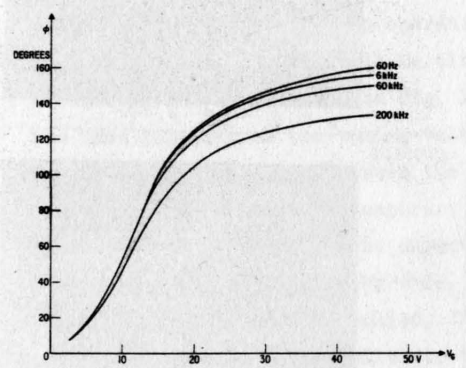


Fig. 11 Phase angle as function of gate control voltage V_G; circuit of Fig. 10, with $R_0 = \infty$. C = 2, 0.02, 0.002, 0.0006 μF, for 60 Hz, 6, 60, 200 kHz, respectively.

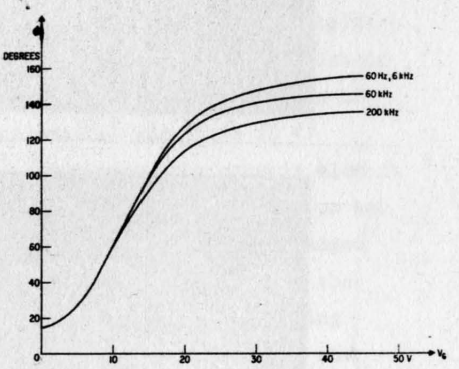
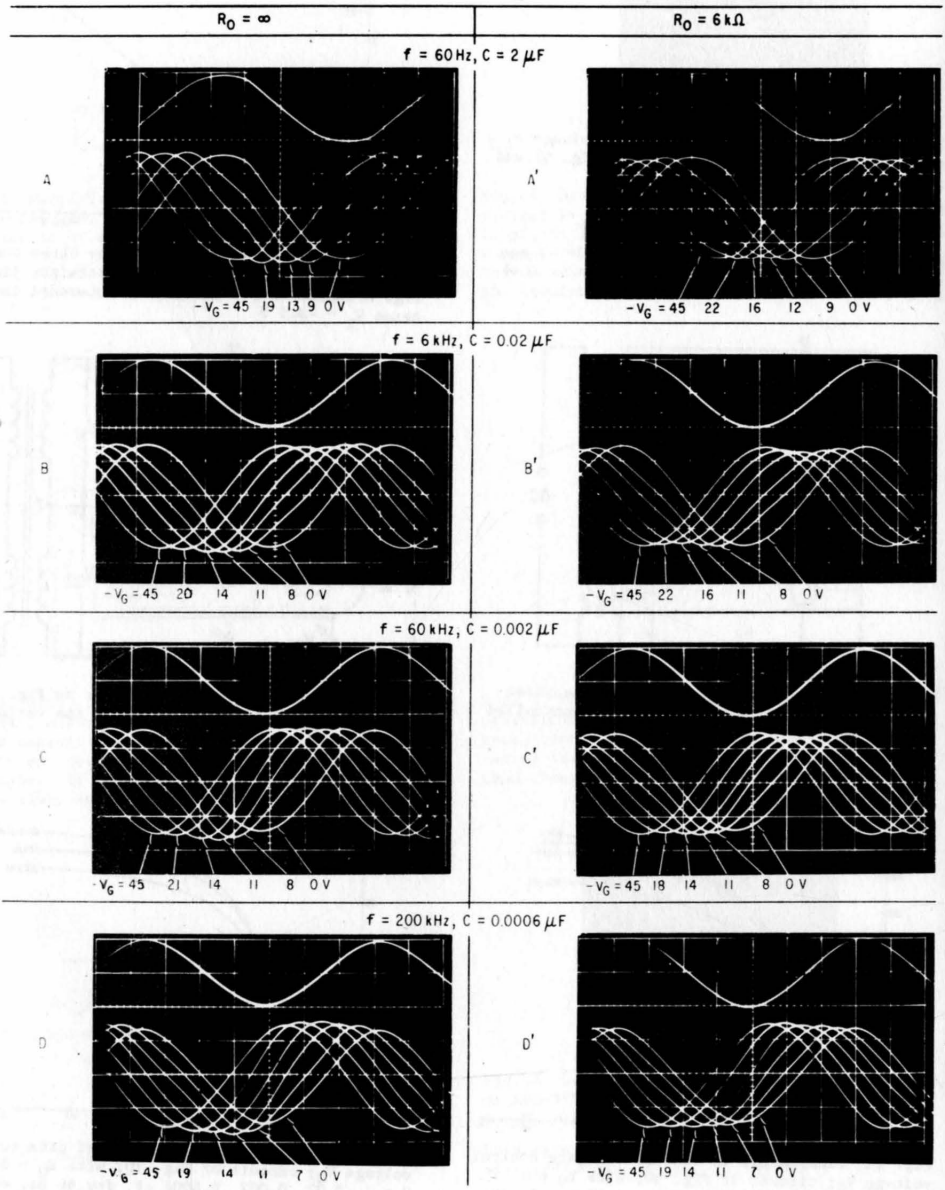


Fig. 12 Phase angle as function of gate control voltage V_G; circuit of Fig. 10, with R_0 = 6 k-ohms. C = 2, 0.02, 0.002, 0.0006 μF, for 60 Hz, 6, 60, 200 kHz, respectively.

Fig. 13 Voltage wave shapes encountered in phase shift circuit of Fig. 10

Upper traces: Secondary voltage V_{S1}, scale approx. 5 V/cm
Lower traces: Output voltage V_O of phase shifter, scale approx. $3\frac{1}{3}$ V/cm



SOLID STATE ELECTROSTATIC CONTROL ELEMENTS OPERATING ON PIEZOELECTRIC PRINCIPLES

by F.L. N-NAGY,Dipl.Ing,CEng,MIEE
Electrical Engineering Department,
University of Salford, England.

1. Introduction

The growing importance of, and further demands for, various new control and computer dynamic components have led us to pioneer a new terrain in which new control elements might be devised. This terrain is the solid state field; new ceramic control components have developed operating on the "inverse piezoelectric effect". Mechanical strain is generated by the charges in the microscopic structure of the ceramic material due to the controlling signal, thus providing a controlled macroscopic action. The interaction of elastic and electrical properties provides a simple and convenient method of obtaining electro-mechanical energy conversion, which is analogous to "motor" operation.

2. Basic Theory

The ceramic materials are isotropic and exhibit piezoelectricity only when they are prepolarised which process is called poling treatment[1]. In order to explain and illustrate the basic mechanisms of the inverse piezoelectric effect for two operating modes of interest to us, a positive and negative d.c. signal will be alternately applied on a ceramic element of dimensions, $\ell \times w \times t$ (see Fig. 1). When a positive d.c. voltage of the same polarity as the poling voltage but smaller magnitude is subsequently introduced between the poling electrodes, the ceramic element experiences a further but temporary expansion in the poling direction and a contraction parallel to the direction of the electrodes. The changed dimensions will be denoted by $\ell-\Delta\ell$, $w+\Delta w$ $t+\Delta t$. Conversely when the negative d.c. voltage is applied, the element contracts in the poling directions and expands parallel to the electrodes. The dimensions now become $\ell+\Delta\ell$, $w-\Delta w$, $t-\Delta t$. When the d.c. voltage has been removed from the electrodes, in either case the element returns to the original poled dimensions. The process is based on thickness and transverse effects; thus the element is called a thickness poled transverse expander.

When the poling electrodes are removed and the element is provided with signal electrodes perpendicular to the poling direction, a d.c.

voltage is applied again and shear takes place around an axis perpendicular to both the poling and signal field in a counterclockwise sense, which can be a parallel or a transverse shear action.

The electrical condition of the ceramics is defined by two vector quantities: the electric field strength E, and the electric displacement D. For high driving electric fields, their relationship shows the same hysteresis characteristic as ferromagnetic materials; hence this D/E dielectric hysteresis phenomenon is called "ferroelectric"[2].

3. Piezoelectric devices

The basic piezoceramic components suitable for an electro-mechanical actuator in a control system are known as bimorph and multimorph elements. A two plate laminar composite variety of piezoelectric bars and one bimorph disc is shown in Fig. 1. All these are manufactured of Pb $(Zn,Ti)O_3$, Lead -Zirconate-Titanate ceramics[3]. A considerable number of ceramic materials exhibit piezoelectric phenomena. However, from a control engineering point of view at present only the polycrystalline PZT compound ceramics are of much interest because their electrical, mechanical and piezo-electric properties seem to be superior to any other ferroelectric materials.[4]

All the bimorph elements[5] in Fig. 1 consist of two plates, electro-statically prepolarised in the thickness direction and cemented together face to face. An excitation signal applied to electrodes on its major surfaces causes opposite longitudinal deformation in the two laminar plates perpendicular to the axis of polarisation, resulting in flexure. Operating in such a mode, the bimorph in a cantilever beam mounting is shown in Fig. 2a. The two plates of the bimorph could also be parallel connected in a longitudinal expander mode. A parallel connected bimorph bar has the advantage that it has a lower operating impedance than the dimensionally equivalent series system and the voltage for equal displacement will be half that for the series connected operation. One disadvantage is that in parallel connection there will be three terminals instead of two. The parallel connected bimorph bar is shown in Fig.2b. From a mechanical point of view the bimorph bender behaves like a spring with an internal stiffness, when it is subjected to an electric field. The stress T_1, i.e. the internal force developed per unit cross-sectional area, produces an effective bending moment, as shown in Figs. 2a and 2b.

For further explanation consider an a.c. driving signal. For the series connection during the first half cycle one half of the voltage

falls across the top plate and is in the direction of the electric displacement. This plate experiences a stress tending to cause a decrease in its length. Similarly, since at the same time the other half of the signal applied to the bottom plate, is of opposite polarity to the internal prepolarisation and develops a mechanical force causing it to increase along its length, the piezoelectric beam will bend upwards. During the following half cycle the process in the plates will be reversed and the beam will bend downwards.

The other type of bimorph is the torsional element. The signal field E_1 is perpendicular to the axis of polarisation in the face plane itself. In addition it can operate either in "face shear mode" or "transverse shear mode" and is torsional about its longitudinal axis.

A flexural and torsional element can be built in one unit and this is the flexural-torsional bimorph, which can operate in these two distinct modes either simultaneously or selectively. The composite effect, viz. bender and face shear mode, is indicated by the dotted line position in Fig.2c, and the relation between polarisation and driving field, in Fig.2d.

All types of bimorph in bender and torsional mode operation can be used in themselves as control actuators as will be outlined in the following sections.

Furthermore the bimorph bars can be assembled into bridge construction for various types of "motor" actions. A working model of a two-phase/one-bridge crystal motor is shown in Fig.3 and its perspective representation in Fig 4. In Fig. 4 "a" denotes the bimorph bars. There is a flexible perspex bridge designated by "b" connecting the two bimorph bender bars together. If they bend one way by equal amounts, the motion of the elbow is horizontal. When they move opposite ways, the motion is vertical.

All types of solid-state control element outlined in this article operate in two modes, viz. in non-resonant condition, having a number of facilities controlling the angular rotation of a driven shaft. Consider first the various types of non-resonant motor operation. The drive of these motors is similar to that for any two phase conventional servomotors. They have two phase signal, one for each bimorph, $0 < \varphi < 180$ degrees apart in time, and $0 < \alpha < 180$ degrees apart in space. Usually φ and α are both 90 degrees. Considering the time and space relation of the driving signal, the resultant motion is

$$f(t, \varphi, \alpha) = e/2 \; e^{j\omega t} + e/2 \; e^{j(\omega t - \varphi)} \cdot e^{j\alpha} = e^{j\omega t}/2 \; (1 + e^{j(\alpha - \varphi)}) \Big|_{\varphi = \alpha = \pi/2} = e^{j\omega t}$$

This is the equation of a circle with an angular speed ω; that is, the driving elbow rotates without load in a circle. When load is applied, e.g. in the present case driving a magnetic ¼ inch tape against a bearing capstan, the elbow of the bridge may rotate in a circle, ellipse, or any pattern with loops depending upon the value of φ and α. The motion could be further influenced easily by superposing harmonics on the driving signal. The ideal superposition of harmonics is when they are zero-phase-locked with the fundamental. This type of driving signal will give an almost straight line motion during a good third of the time. The speed could be controlled by varying either the amplitude or frequency of the driving signals or controlling the time phase angle φ. The amplitude and phase control could operate at resonance, hence the name piezo-resonant motor. Their torque-speed characteristic in both nonresonant and resonant mode has great variety regarding quite a number of controlling facilities to influence the pattern of motion. Fig. 5 shows a complete two-phase/two-bridge crystal motor, which consists of two bridge units, two opposite parts, composed of four bimorph elements, so that the tape is punched from both sides. When one pair releases the tape, the other pair takes over. A further development of a crystal motor with a more conventional operation is shown with a rotating shaft in Fig. 6. This type of motor could also be developed in three phase or multiphase construction. All types of crystal motors which have been presented here, have further consideration facilities, viz building more units of one type of basic motors in mechanically parallel cascaded connection. By this method the resultant torque on the rotating shaft can be increased quite considerably. The only reason for using the tape drive application was to provide some information about single unit motor performance.

For torque-speed characteristics measurement a new device had to be developed, called Servo-Torque-Balance, which is able to measure even very small torque in the range of 1 g cm. The maximum torque for a single unit is in the range 5-25 g cm, depending on the pattern of motion. But for the rotating version of two-phase/two-bridge type motor, the maximum expected torque is around 20-100 cm and for parallel cascade connection the torque could be few hundred g cm.

Furthermore two constructions of controllable pumps are shown which

can be applied with either pneumatic or fluid operation. In Fig.7 the actuators are two cantilever mounted bimorph bars operating the membrane in the pump. Increase of movement has been introduced by a second order level of ratio 1:4. In Fig. 8 the membranes are replaced by two bimorph discs operating in piezoresonant mode.

4. Bimorph actuator in control loop

In the following a novel application of the bimorph is outlined as an electromechanical transducer controlling the alignment of a magnetic head relative to a multitrack computer store tape. The bimorph acts as a conventional controlled vibrator in a servo-system.[6-13]

The magnetic head is supported in an assembly and coupled to piezo-electric bimorph elements which serve as actuators in servo loop designed to compensate skew, wow and flutter tendencies in instrumentation tape recording devices.

An assembly as shown in Figs. 9 and 10 was used in an angular position servo-mechanism to control the magnetic head for a computer store. The moving parts of the assembly, consisting of two piezo-electric bimorph actuators, a 16 track magnetic head and mechanical links between them, formed a fairly heavy device, the mass of which was represented by a considerable inductance in the equivalent circuit. Consequently, it reduced the first resonant frequency of the bimorph, and with this of course the controllable frequency range.

For calculation of the resonant frequency, the distributed mass of the vibrating cantilevel bimorph may be replaced by a point mass of approximately one quarter its magnitude. This value can be derived from the conservation of energy in the bender bar.[13] An appreciable reduction of the effective mass of the magnetic head was also necessary. Due to the geometrical arrangement the mass of the rotationally vibrating head could be replaced by an effective mass of one third its magnitude. Then further reduction was obtained by means of a second-order level. By the second method the effective mass of the magnetic head was reduced quadratically in the ratio of the arms, but at the expense of a proportional decrease of movement. Altogether the effective masses of the crystal and magnetic head were reduced from 7.1 to 1.64 g and 30 to 2.5 g respectively. The first resonance decreased from 290 to 118 Hz.

A working arrangement of a crystal angular-position control system is shown in Fig. 11, partly in perspective and partly in block-diagrammatic representation. In the diagram, 3 represents one inch wide tape, 2-15

shows that part of the tape which is used for receipt of information of fourteen parallel tracks. Another two trakcs, 1 and 16, are indicated. On these are recorded reference square-wave signals of constant frequency (25KHz) and constant amplitude, shifted 90 degree from each other. They permit detection of the misalignment of the magnetic head, that is, the skew. If the tape skews, the reference signals have a phase difference other than 90°degree. The played-back reference signals from tracks 1 and 16 are amplified by limiter amplifiers 4 and 5 with about 100 dB overall amplification to obtain adequate magnitude and rise time for the reproduced square-wave signals. The reproduced signals are then fed to a phase-error discriminator, 6, which can detect both the sign and magnitude of the error in the form of a difference square-wave signal. In the following two stages this is integrated and amplified in a d.c. amplifier of modulated version, 7. The output of the amplifier drives the bimorph transducer, 8, which is fixed at one end in a block, 9. The other end is coupled to the magnetic head, 10, by the rigid member, 11, and is free to move in the direction of arrow 12. The head itself is mounted on a shaft, 13, in a fixed bearing, 14. The free end of the crystal rotates the magnetic head under the controlling error signal, so that the gap alignment is perpendicular to the tape movement. Thus the phase error between the two reference signals, and with it the skew, is reduced in all fourteen tracks.

For wow and flutter compensation, one of the reference signals, e.g. track 1, is amplified by 5, and fed to the frequency discriminator, 17. An oscillator, 18 feeds a reference signal of the same frequency as the square-wave signal recorded on track 1. The result is the error signal which varies according to tape speed variations. The error is amplified by a d.c. amplifier, 19 and fed to a bimorph, 20. The wow-flutter servo actuator is fixed at one end, 21, and the other end is free to move in the direction of arrows 22, about the pivot 23 in a fixed bearing 24 mounted in a rotating block. Thus the magnetic head is controlled by the crystal actuator over an arc which is substantially parallel to the direction of motion of the tape and reduces the wow and flutter irregularities in the tape speed.

5 Derivation of transfer function and stability consideration.

The transfer function of such a basic piezo-electric device, i.e. a single bimorph in motor application, might be derived on the basis of an electro-mechanical transformer as its equivalent circuit and the use of

pure mechanical bending theory.

The more precise transfer response can be derived by applying D'Alembert's principle to obtain the equation of motion of a bimorph bender bar. When its two dimensions are of the same order, as for example our case here, that is, the width becomes comparable with the length and only the thickness is relatively small, the equation of motion of this structure is given by

$$\rho \frac{\partial^2 x}{\partial t^2} + Y_{11}^E K \frac{\partial^4 x}{\partial t^4} - \rho K^2 \frac{\partial^4 x}{\partial x^2 \partial t^2} = 0 \qquad \ldots (1)$$

where Y_{11}^E, K and ρ are the Young's modulus, radius of gyration of the section and specific density of the crystal bar respectively. The last term in equ. (1) represents the force reaction of the rotary inertia, which results from the added kinetic energy of the section rotating about an axis located normal to the central line of the bar. When dealing with electrostatic controller we have also to incorporate the mechanical electrical and piezoelectric relations. In a crystal producing a single mode of motion, for example in our case only the longitudinal mode, i.e. expansional and contractional when the driving voltage V is applied, the equations for electrical and piezoelectric relations are

$$T_1 = Y_{11}^E S_1 - \frac{Y_{11}^E d_{31}}{\epsilon^T} D_3$$

$$E_3 = -\frac{Y_{11}^E d_{31}}{\epsilon^T} S_1 + \frac{1}{\epsilon^T} D_3$$

where S = strain; T = stress; D = electrical displacement; E = electric field strength applied by voltage V; d_{31} = piezoelectric strain constant; ϵ^T = dielectric permittivity at T = 0; Y_{11}^E = Young's modulus at E = 0

Considering the bimorph as a four-terminal electro-mechanical network, the mathematical presentation of the transfer function is implemented using four pole theory[8] and incorporating the equations (1) and (2) with certain idealisations. The idealisations and simplifying assumptions are given elsewhere.[6]

The system consists of an amplifier, cascade compensating lead-lag circuit and the crystal transducer assembly, as shown in Fig. 12 Assuming small

movement and approximately linear operation, the most simplified open-loop transfer function of the system with the incorporated bimorph assembly is

$$G(s) = K_a \frac{1 + sT_1}{1 + sT_3} \frac{1 + s\,T_2}{1 + s\,T_4} \frac{k^2/n\,C_o}{s^2T_o^{\,2} + 2s\,\zeta_o T_o + 1} \qquad \ldots(3)$$

where K_a = gain of amplifier, T_1, T_2, T_3, T_4 = time constants of lead-lag compensating circuit, k = coupling factor, n = piezo-electromechanical transfer ratio, C_o = static capacitance of bimorph, T_o = electromechanical time constant of bimorph, and ζ_o = damping factor.

For stability determination, assuming that the signal modulated amplifier has a small output achieved by applying large envelope feedback, we apply the general Nyquist stability criterion. Stability of operation even with a large gain amplifier, is ensured by the use of compensating circuit with properly chosen time constants. In the case of larger movements, up to the mechanical limit of the bimorph, the system becomes nonlinear with hysteresis and butterfly characteristics. The operation under these conditions is under investigation in order to derive the criteria for optimum responses.

For control applications, ceramics of either Types PZT4, PZT5, and PZT-5H could be chosen for their low elastic Q factor, which makes them more suitable for servo-transducer application. The table shows the significant coefficients of two types of Lead-Zirconate-Titanate ceramics. Recently other types of ceramics suitable for this application have become available, and these are also listed in the table mentioned above. The coefficients are required for the calculation of such bimorph parameters as effective inertial mass (14), compliance (15), natural frequency(16,17), etc., during 'inverse' piezo-electric operation. The parameters in the table are needed in order to compute numerically the coefficients in the transfer function of the bimorph.

Electrostatically controlled ceramics as actuators in servo-systems have some intrinsic advantages. They are much small than an equivalent electro-magnet device and are immune to electromagnetic disturbances at the magnetic head. Furthermore, in the application described above, the computer-store tape speed and tension, motor inertia, eccentricity of the capstan, etc., need not be so accurately designed as they would have to be without bimorph control. The effect of irregularities in the motion of the tape is

reduced by a relatively cheap servo-system controlling the magnetic head.

The crystal-controlled skew compensation for the computer magnetic store has produced an average improvement of 20-30 dB in the subresonant frequency range, 20 - 100 Hz.

6. Further things to come

Ceramic crystals have already revolutionised the practice of both "direct" and "inverse" electromechanical transduction. There is considerable room for exploratory work on ceramics as control actuators especially when the whole movement is only in the range of up to a few hundred microns in magnitude.

It is worth mentioning that other researches and developments are now in progress for the application of multimorph ceramic elements and multimorph control actuating assemblies. The multimorph elements for flexure, torsional or both together in themselves are also promising servo actuators. They are somewhat more stable than the two plate bimorph versions.

Every year produces further ceramic compounds with improvements in the stability of their constants, which makes them still more suitable for use as electrostatic servo controllers. In general we must accept that no single ceramic compound at present available is capable of meeting all the requirements for control application. Therefore the directions of research will be turned to synthesising new ferroelectric compounds that will satisfy as many as possible of the requirements on control applications.

7. Acknowledgements

I am indebted to Professor D Gabor for his initiative suggestions and furthermore to Professor J H Calderwood, Chairman of the Electrical Engineering Department at the University of Salford, for his support and provision of all the necessary facilities in r. & d. on crystals in servo system applications.

References

1. Deri, M.: "Ferroelectric Ceramics," Maclaren and Sons Ltd. London 1966.

2. Brown, C.S. and others: "Piezo-electric Materials". Proceedings of IEE, Vol. 109, Part B, Number 43, January, 1962.

3. Crawford, A.E.: "Lead Zirconate Piezo-electric Ceramics", Brush Crystal Co Ltd. Presentation, Southampton.

4. Fischbacher, R.E.: "Piezo-electric and Magneto-strictive Elements", Control, September 1959.

5. Mason, W.P.: "Piezo-electric Crystals and their Application to Ultrasonics, Van Nostrand, 1950.

6. N-Nagy, F.L.: "A novel electrostatic actuator in a direct feedback system, "Control 11(108), June 1967, pp.293-296.

7. N-Nagy, F.L.: "Piezo-electric element as Novel Electrostatic Controller in a closed Loop Servo System", Conference on Servocomponents sponsored by the IEE Control and Automation Division, 21st-23rd November, 1967, Conference Publication No. 37, pp 165-173.

8. N-Nagy, F.L.: "The cantilever beam mounted bimorph crystal as a four-pole element" Internal Report at University of Salford.

9. N-Nagy, F.L.: "Compensation of irregularities in motion of a recording medium", British Patent No. 925-603, Specification published May 8, 1963.

10. N-Nagy F.L.: "Report for Wolfson Foundation on "Solid/State Control Components and Devices etc., Operating on Ferroelectric Principles", University of Salford, July 1968.

11. N-Nagy, F.L. : "Ujszeru elektrosztatikus beavatkozo szerv merev visszacsatolasu rendszerben." Meres es Automatika, XVI. evfolyam, 1968. 6. szam, pp. 263-265.

12. N-Nagy F.L.:
Пьезоэлектрический преобразователь в качестве исполнительного элемента системы управления, РЕГУЛЯТОРЫ И СЕРВОМЕХАНИЗМЫ, 156, УДК 62-528, Экспресс Информача, 1967, №31, стр. 34 - 41.

13. N-Nagy, F.L.: "Procede et appareil pour compenser less irregularities de mouvment d'un support d'enregistrement"
French Patent No. 1.309.568. P.V. no 882.552.

14. Timoshenko, S. and Young, C.H.: "Vibration Problems in Engineering", Van Nostrand, New York, 1958.

15. Mason, W.P.: "Electromechanical Transducer and Wave Filters", Second Edition Van Nostrand, 1948.

16. Olson, H.F.: "Acoustical Engineering, Van Nostrand, New York, 1960.

17. Orwell, R.F.J.: "Piezo-electric Flexure Transducer Elements", Ultrasonics, Jan-March, 1963.

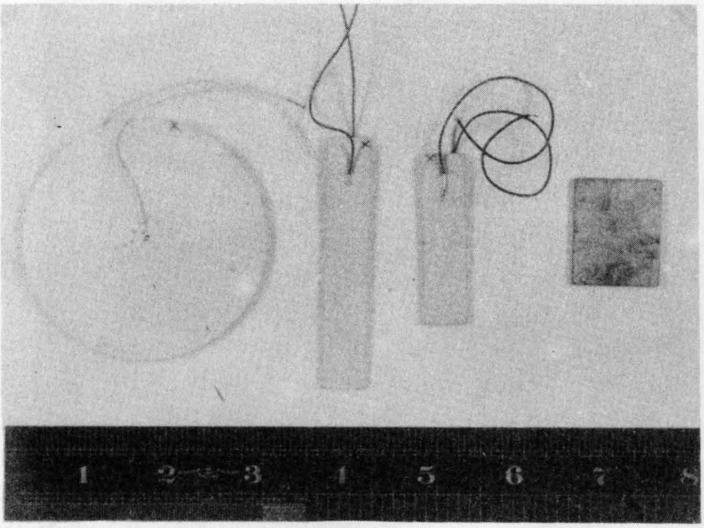


Fig.1 Two plate laminar piezoelectric bars. Sizes are from right to left: (i) 2.5 x 2.5 x 0.15cm; (ii) 5 x 1.25 x 0.15 cm; (iii) 7.3 x 1.25 x 0.15cm; (iv) diameter 7.3 cm, thickness 2 mm.

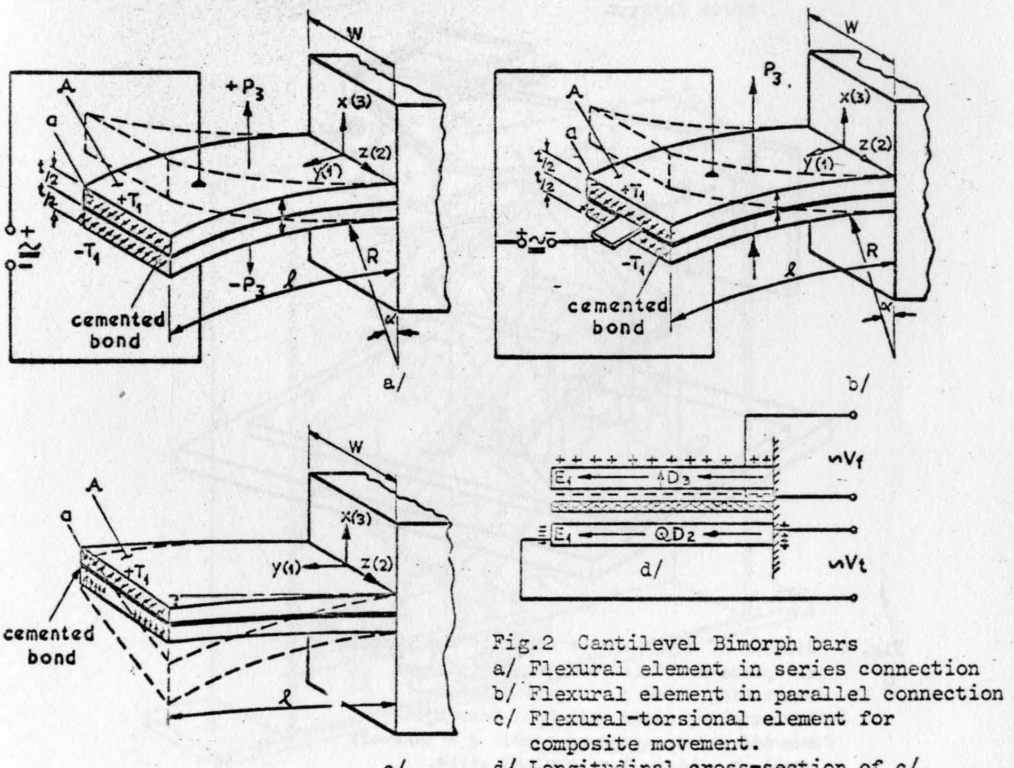


Fig.2 Cantilevel Bimorph bars
a/ Flexural element in series connection
b/ Flexural element in parallel connection
c/ Flexural-torsional element for composite movement.
d/ Longitudinal cross-section of c/

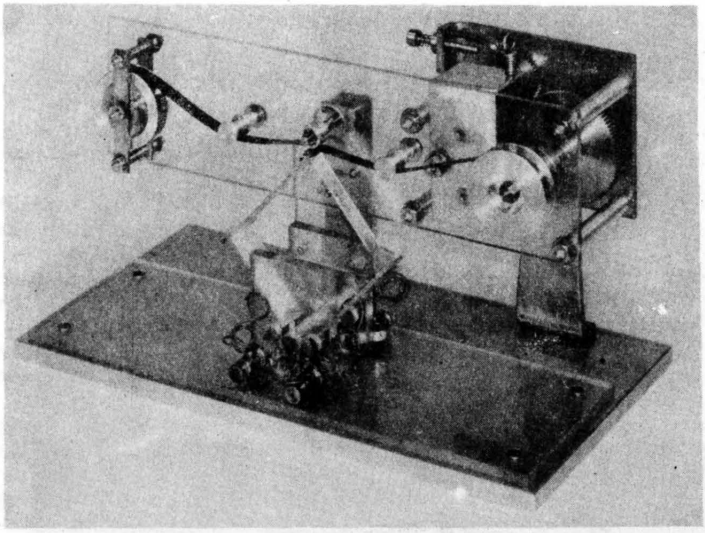


Fig. 3 Piezo-resonant motor incorporating two bimorph cantilever flexural elements - see fig. 4 for key

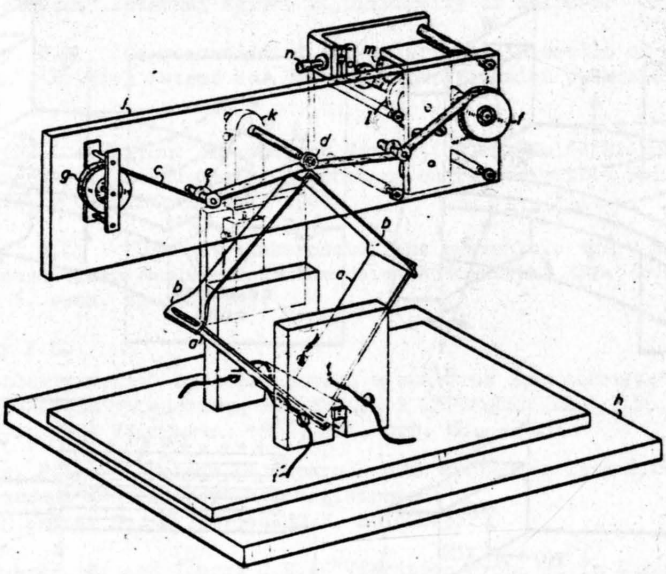


Fig. 4 Diagram of cryatl motor assembly and symbols
a - piezo-electric element. b - perspex connecting beam. c - magnetic tape.
d - rotating idler with ball bearing.
e - idlers. f - take-up spool. g - take-off spool. h - base. i - terminal wires.
j - back panel. k - adjusting rod. l - take up spool spring motor. m - reduction gear.
n - ratchet release button. o - elbow

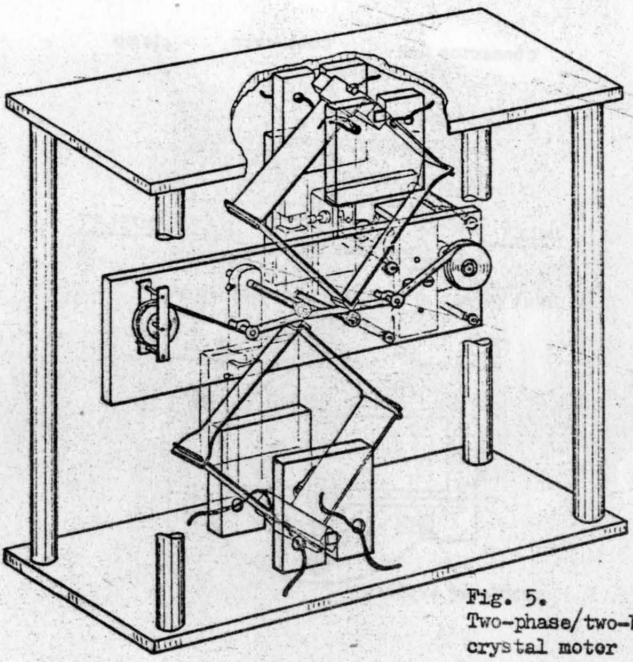


Fig. 5.
Two-phase/two-bridge
crystal motor

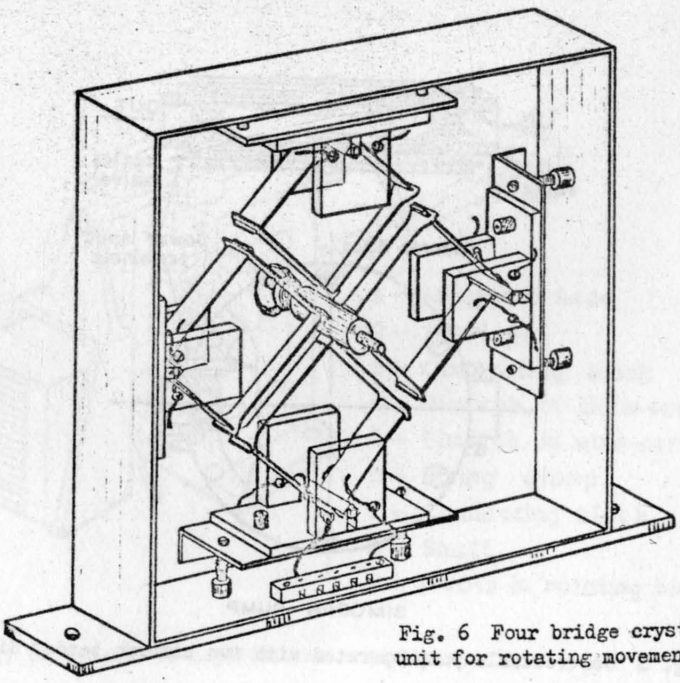


Fig. 6 Four bridge crystal
unit for rotating movement

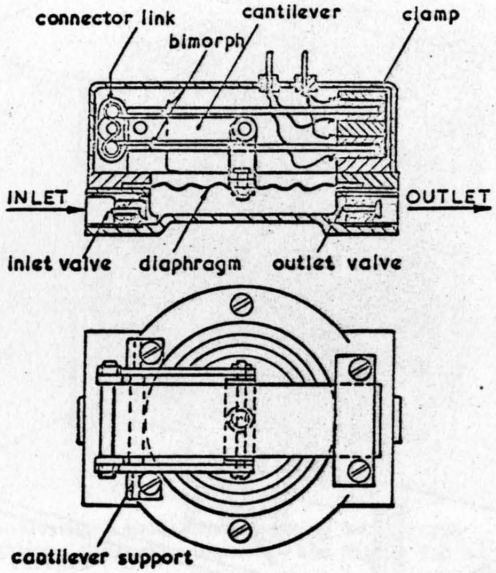


Fig. 7 Controllable pump operated by two cantilever mounted bimorph bars

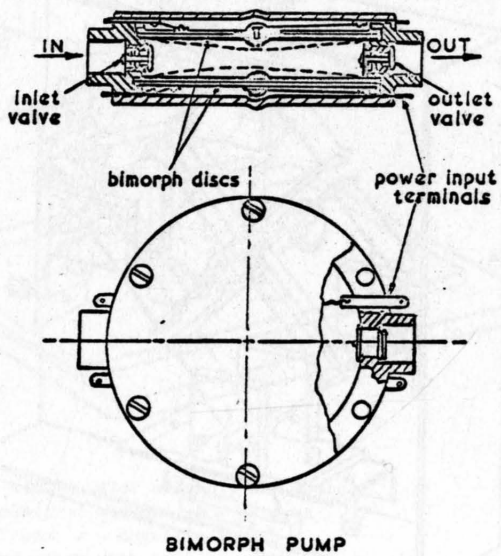


Fig. 8 Controllable pump operated with two bimorph bender discs

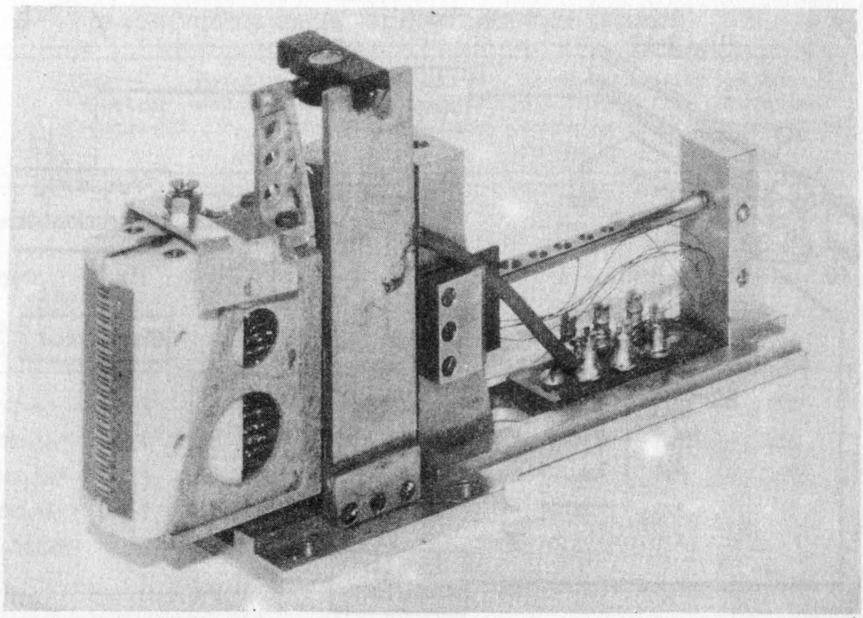


Fig. 9 16 track magnetic head assembly incorporating bimorph as actuator - see figure 10 for key

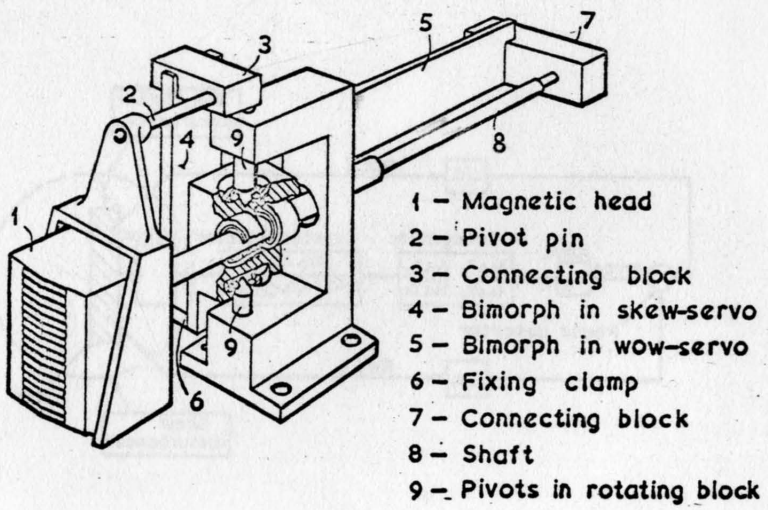


Fig. 10 Diagram of magnetic head assembly

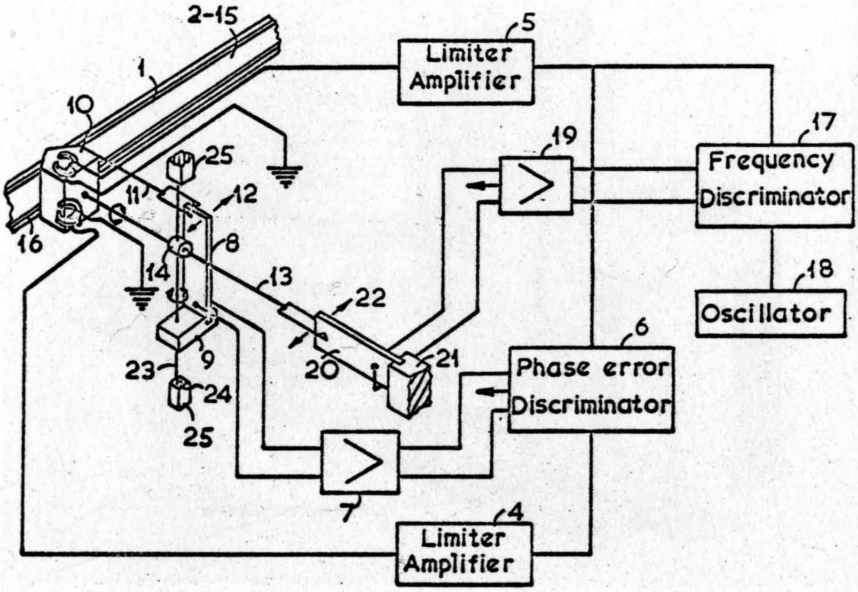


Fig. 11 Operational diagram of crystal angular-position control system

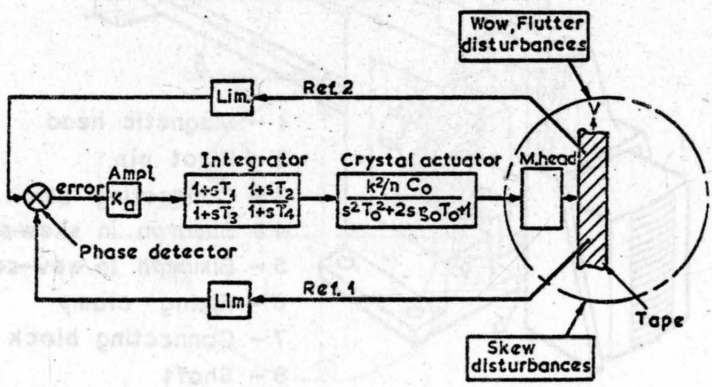


Fig. 12 Block diagram of servo system with bimorph bar actuator.

TABLE

OF COEFFICIENTS OF LEAD-ZIRCONATE-TITANATE CERAMICS

	Piezo-electric contstant	Young's modulus	Elastic quality factor	Relative dielectric constant	Piezo-electric coupling factor	Density	Maximum Operating Temperature
Symbols	d_{31}	Y^E_{11}	Q	ε^T	K_{31}	ρ	T
Units	m/V	N/m^2	-	-	%	kg/m^3	^oC
Multi-plier	10^{-12}	10^{10}	-	-	10^2	10^3	-
PZT-4*	-123	8.15	500	1300	0.334	7.5	200
PZT-5A*	-271	6.2	75	1700	0.344	7.75	250
PZT-5H*	-171	6.15	65	3400	0.388	7.5	130
HST-41**	-157	7	70	1800	0.35	7.4	-
G-1512**	-232	6.3	70	2600	0.37	7.4	-

*Brush-Clevite Ltd., **Gulton Industries Inc.

PRECISE HIGH-SPEED ABSOLUTE POSITION CONTROL USING A MULTI-TRACK OPTICAL GRATING

A. Russell
National Engineering Laboratory
East Kilbride, Glasgow, Scotland

1. TRANSDUCER REQUIREMENTS

In the field of precise high-speed digital positioning such as is encountered in on-line digital plotting and integrated circuit mask manufacture, there is a need for a transducer and a measuring system which will fulfil as many as possible of the following requirements.

(a) Give accurate and precise measurement.

(b) Traverse at high rates without loss of information.

(c) Have minimum friction.

(d) Be immune to electrical interference.

(e) Have minimum inertia.

(f) Have a minimum of electronic complications.

(g) Have datum-shift facility.

Owing to the high frequency of operation it is not considered necessary to have a digital read-out from such a system. If it is essential however this read-out can be obtained from the digital input data with the system at balance.

Previous work at National Engineering Laboratory in extracting absolute digital displacement information from multi-track optical gratings resulted in the development of the NEL absolute displacement measuring system[1,2] and in a novel point-to-point positioning system[3]. The absolute behaviour of the multi-track optical grating fulfils requirements (a), (b), (c) and (d) and, by using plastic instead of glass for the grating base, (e) can be satisfied in some measure.

2. MULTI-TRACK OPTICAL GRATING

The accuracy and precision of optical gratings has been well established[4] and requirements (b) and (d) can be satisfied if positional information can be extracted in absolute form. (Absolute is defined as existing by itself without relation to other things.) If coded patterns are used resolution is lost owing to the impossibility of interpolating the digital 'bit' size. However, although individual cycles of the grating waveform cannot be separately identified, each cycle can be divided into many parts by the technique illustrated in Fig. 1, and each intermediate waveform can be given an absolute identity. Where one cycle of grating waveform represents the full stroke of a system each intermediate waveform will represent the most significant digit of the input data.

In Fig. 1 the sine and cosine basic grating waveforms and the inverted sine ($\overline{\sin}$) and basic cosine waveforms are applied to equal value resistor pairs. Intermediate waveforms are generated at the resistor junctions. These cross the zero-volt level at 45 and 225 degrees, and 135 and 315 degrees respectively. The amplitude of the R1/R2 intermediate waveform (V) will be

$$V = \cos\theta + (\sin\theta - \cos\theta) \times \frac{R_2}{R_1 + R_2} \tag{1}$$

and the amplitude of the R3/R4 intermediate waveform (W)

$$W = \cos\theta + (\overline{\sin}\ \theta - \cos\theta) \times \frac{R_4}{R_3 + R_4}. \tag{2}$$

A suitable potential divider can be chosen which will produce any desired number of intermediate waveforms which cross the zero-volts level at the desired intervals. The potential divider consists of a chain of resistors which will yield at their junctions the potential division (P) of sin/cos voltages

$$P_{s/c} = \frac{R_1}{R_1 + R_2} = \frac{\sin\theta}{\sin\theta - \cos\theta} \tag{3}$$

and for cos/$\overline{\sin}$ voltages

$$P_{c/s} = \frac{R_4}{R_4 + R_5} = \frac{\cos\theta}{\cos\theta - \overline{\sin}\ \theta}. \tag{4}$$

From grating tracks of 4 lines or less per mm (100 lines/in), however, serrisoidal (linear) signals are emitted and the potential divider can be completely linear.

It will be observed from Fig. 1 that, if either of the intermediate waveforms were used as a control error waveform, it would have a tolerance of $\pm 180^{\circ}$ or $\pm\lambda/2$ before running into phase reversal error. For instance, if it was desired to lock-on to position 5 (sin or $\overline{\sin}$ waveform), the error signal due to variations from the balance point would be of correct polarity till position 1 on either side, beyond which the phase-reversal error would cause a new lock position at position 5 $\pm\lambda$.

$$\text{Phase-reversal error tolerance} = \text{tol}_{\beta} = \pm\frac{\lambda}{2}. \tag{5}$$

If Fig. 1 represents one cycle of information from the coarsest track of a multi-track grating and each zero-voltage point represents a most significant digit of the input data, then it will be seen that by including a second track eight times finer than the first and similarly dividing this into eight discrete parts $\lambda_1/8$ of the first track will be equal to λ_2 of the second track and the $\lambda_2/8$ points will represent the second most significant digit of the input data. Thus a series of discrete points can be

provided representing, in this case, digits in octal of multi-digit input data.

If a position control servo system is balanced by the data-selected intermediate waveforms from each track in turn, starting with the most significant, an absolute analogue positioning system which will fulfil requirements (a) to (f) is feasible.

3. PRACTICAL TOLERANCES

For a more practical decimal system with, say, a total stroke of 1 cm, it is essential to control the transducer from any instantaneous position within the stroke to any other position dictated by the input data. From equation (5) it is obvious that the total stroke cannot be greater than $\lambda/2$ in order to keep within the phase-reversal tolerance (tol_β) of $\pm\lambda/2$.

Thus, for λ = 2 cm and a stroke of 1 cm, a suitable 180° potentiometer for a grating cycle of 2 cm may be as shown in Fig. 2(a). The resultant intermediate control waveforms (serrisoids for simplicity) are shown in Fig. 3. The second most significant track has λ = 1 mm and a full 360° potentiometer, as shown in Fig. 2(b), is required for continuity. Assuming the data input to be 8.5 mm, one must consider the accuracy required at point 'X_n' before control can be switched over to the second track at point 'X_{n-1}'.

If one designates the 1 cm stroke of the coarse track 'λ_n' and each sub-division unit 'δ_n' (1 mm), and in the fine track one cycle and one sub-division 'λ_{n-1}' and 'δ_{n-1}' respectively, then the main tolerance will be the phase-reversal tolerance (5) which will now be

$$tol_\beta = \pm\frac{\lambda_{n-1}}{2}. \qquad (6)$$

Since there is no particular phase relationship between the grating tracks, the '0' waveform chosen on the 'n - 1' track will be the one nearest coincidence with the 'n' track intermediate waveforms, both at zero-volts condition. This, however, could give a worst-case error of

$$\text{grating phase displacement tolerance} = tol_\theta = \frac{\delta_{n-1}}{2}. \qquad (7)$$

From tolerances (6) and (7) must be subtracted a selection error tolerance. If the 'X_n' point represents steps of, say, units of 'λ_{n-1}' (as shown), then the 'δ_{n-1}' units selection could vary from the 'X_n' point by $9/10 \times \lambda_{n-1}$. This error is much greater than tolerance (6) above so the 'X_n' point should represent steps $\leq$ 'δ_{n-1}'. The 'δ_{n-1}' units selection could then vary from the 'X_n' point by only $9/10 \times \delta_{n-1}$, within tolerance (6).

In any case the selection error tolerance = $tol_{sel} = 9/10 \times y$, where y = the unit steps of 'X_n'. By the inclusion of an extra resistor, R_{tol},

in the potential divider this tolerance can be expressed as

$$\text{selection error tolerance} = \text{tol}_{sel} = \pm\frac{y}{2}. \tag{8}$$

Combining equations (6), (7) and (8) shows that the 'X_n' point must be accurate to

$$\text{total tolerance} = \text{tol}_{tot} = \pm\left(\frac{\lambda_{n-1}}{2} - \frac{\delta_{n-1}}{2} - \frac{y}{2}\right). \tag{9}$$

In the practical example of Fig. 3 one can substitute values for part of equation (9)

$$\text{(part) tol}_{tot} = \pm\left(\frac{\lambda_{n-1}}{2} - \frac{\delta_{n-1}}{2}\right)$$

$$= \pm\left(\frac{1 \text{ mm}}{2} - \frac{0.1 \text{ mm}}{2}\right) = \pm 0.45 \text{ mm}.$$

Clearly the 'X_n' point will have to select units of 0.1 mm to an accuracy of ±0.45 mm. This can be done with a Kelvin Varley divider where the most significant digit sets the two-pole switches and the second most significant digit selects the appropriate one-pole switch. If a very conservative potential division accuracy of ±1 per cent of λ_n is assumed the accuracy at 'X_n' will be ±0.1 mm. As much better potential division accuracies are possible one must consider larger amounts of division to exploit the system fully.

4. SYSTEM '÷100'

Fig. 4 shows an arrangement for a '÷100' system. This has a two-track grating, coarse track = ½ cycle/cm, fine track = 1 cycle/0.1 mm (100 cycles/cm) and resolves to 1 part in 10^4 (1μ).

A double Kelvin Varley bridge is used in the coarse potentiometer so that 'X_n' output will select units of 0.01 mm (= y). The total tolerance of this arrangement becomes (9).

$$\text{tol}_{tot} = \pm\left(\frac{\lambda_{n-1}}{2} - \frac{\delta_{n-1}}{2} - \frac{y}{2}\right)$$

$$= \pm\left(\frac{0.1 \text{ mm}}{2} - \frac{0.01 \text{ mm}}{2} - \frac{0.01 \text{ mm}}{2}\right)$$

$$= \pm(0.05 - 0.005 - 0.005 \text{ mm}) = \pm 0.04 \text{ mm}.$$

This, for a total stroke of 1 cm, is a potential division accuracy of

$$\frac{0.04 \text{ mm}}{1 \text{ cm}} \times \frac{100}{1} = \pm 0.4 \text{ per cent}.$$

The switch positions in Fig. 4 represent an input signal of 4.445 mm. It should be noted that, owing to the insertion of R_{tol} in the '$\delta\delta_n$' potentiometer, the 'X_n' value will always be that dictated by the first three digits of the input data plus an extra 0.005 mm for tolerance purposes.

This ±0.4 per cent potential division accuracy can be achieved if care is taken in grating signal amplification to preserve waveform and stability; only very close tolerance resistors for the initial potentiometers in both channels and stable driver servo-amplifiers should be used.

An interesting '÷60' system can be used for angular indexing applications where the coarse track consists of 360 cs/rev, this '÷60' (in Kelvin Varley steps of ÷6, ÷10 (÷6)) provides an 'X_n' of units of 10 (÷5) seconds, while a second track of 21 600 cs/rev is subsequently divided into steps of one second by the same method (÷6, ÷10). The potentiometric division accuracy required is ±0.55 per cent.

ACKNOWLEDGEMENTS

This paper is published by permission of the Director of the National Engineering Laboratory, Ministry of Technology. It is British Crown copyright and is reproduced by permission of the Controller of Her Britannic Majesty's Stationery Office.

REFERENCES

1. RUSSELL, A. An absolute digital measuring system using an optical grating and shaft encoder. Instrum. Rev., 1966, 13, 174-175.
2. RUSSELL, A. Technical aspects of the NEL absolute grating measuring system. NEL Report (to be published). East Kilbride, Glasgow: National Engineering Laboratory.
3. RUSSELL, A. Recent application of the optical grating in displacement measurement and control. I.E.E. Conf. on servo components, 1967, Conf. Publication No 37. London: Institution of Electrical Engineers, 1967.
4. BURCH, J. M. The metrological application of diffraction gratings. In Wolf, E. (Ed). Progress in optics, Vol. II, 75-108. New York: John Wiley & Son, 1963.
5. SHEPHERD, A. T. The use of moiré fringe techniques. Instrum. Rev., 14, 1967, 185-186.

APPENDIX

1. THE TRANSDUCER

The relative position of two gratings of similar pitch causes the formation of a moiré fringe pattern. Recent developments[5] have made possible accurate transducers on the principle of a short 'index' grating and a long 'measuring' grating mounted adjacent to each other on the two members, the relative movement of which is to be measured. Modern methods of master grating manufacture can produce gratings of high accuracy and durability; linearities approaching one part in 10^6 have been achieved.

Serrisoids are produced from coarse gratings (which act as a linear shutter) while fine gratings can be adjusted to produce sine waveforms at the photocells with a fair tolerance in 'gap' spacing between grating and index.

As the described positioning system is d.c. throughout some care must be given to d.c. stability. In the fine track shown in Fig. 6 the four indices are arranged in 'space' quadrature. The four photocells are connected into two 'push-pull' pairs, $0-180^\circ$ and $90-270^\circ$. This arrangement produces an increase in amplitude of the sin and cos waveforms while cancelling any d.c. shift caused by variation in lamp brightness; e.g. when a pair of cells produces zero-volts both cells 'see' equal light and produce equal and opposite voltages, this condition will hold good over a wide range of lamp brightness. Variations in waveform amplitude are unimportant as this will not affect the zero-volts position of the intermediate waveforms. Matched silicon photo-voltaic diodes should be used.

In the coarse tracks shown in Fig. 6 the index slits are in-line while the four tracks are in 'space' quadrature (this reduces the size of the index reading head) but the cells are again connected as for the fine track.

The light source should be a lamp and collimating lens with the filament parallel to the fine grating lines. The Frésnel gap between grating/index for 100 lines/cm density is not critical, but a slight 'tilt' between the grating and index lines may be necessary to achieve sine waveforms. In a system ten times coarser (stroke = 10 cm, resolution = 10μ) serrisoids will be obtained from both tracks.

High stability amplifiers are necessary for signal amplification and unity gain amplifiers for insertion at the Kelvin Varley switch points to avoid load variations in the potentiometer.

Resistors of at least 0.1 per cent accuracy are required in the initial potentiometers, but this accuracy can be reduced for further potentiometers in the Kelvin Varley chain. Thermal stability is not important providing no thermal gradients are allowed to develop.

2. DIGIT SHIFT

The system operates on positive integers on both axes but, where datum shift is required for centre zero co-ordinates etc., this is best done by inserting adders between the input data and the intermediate waveform selector (Fig. 5) which leaves the actual system unchanged. The adders can be used to match any input data values to any particular machine datum within the range.

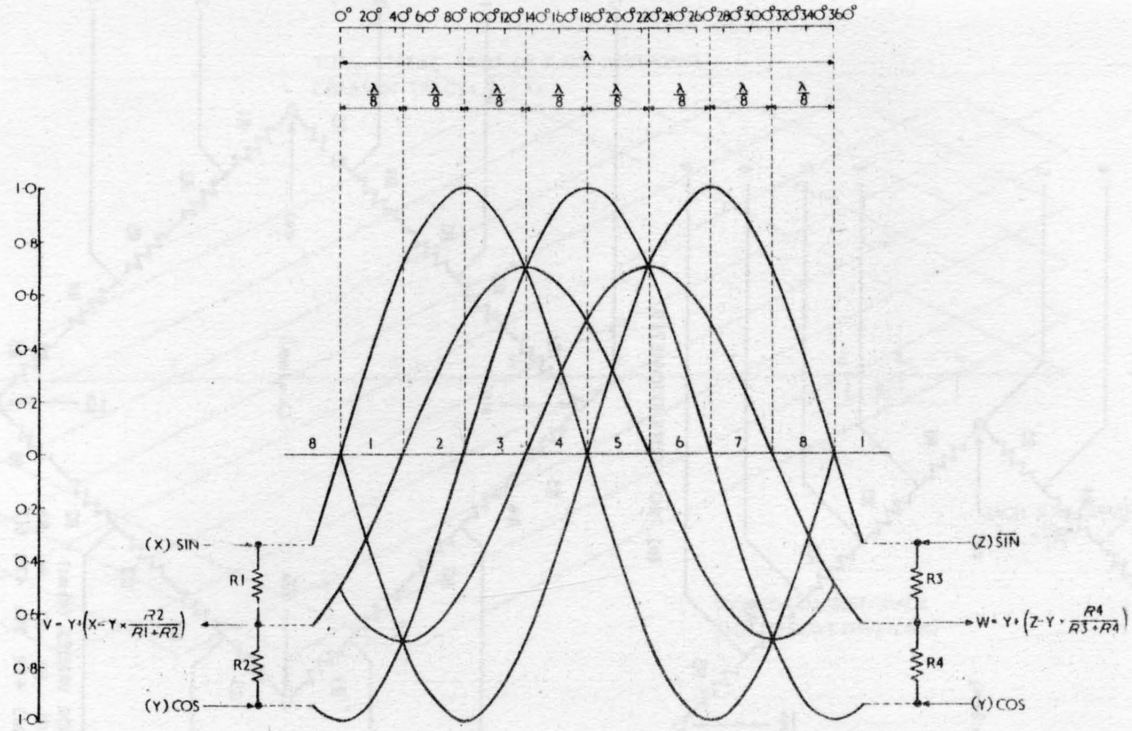
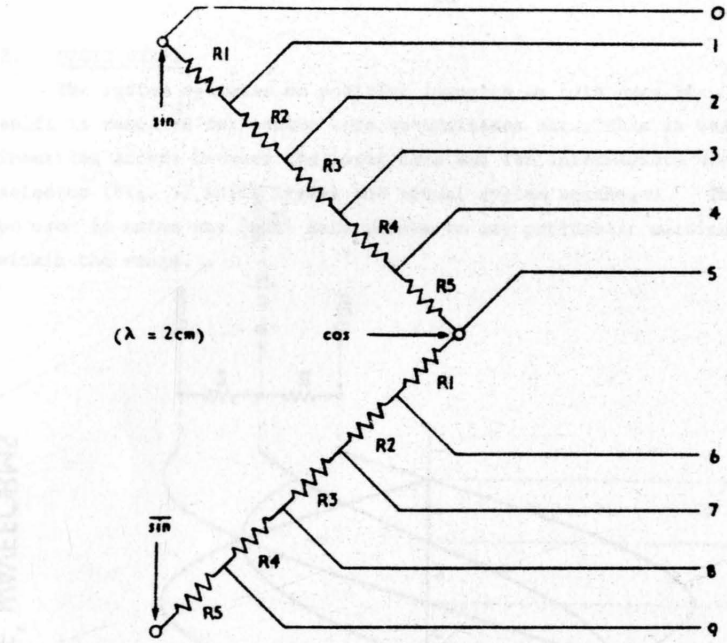
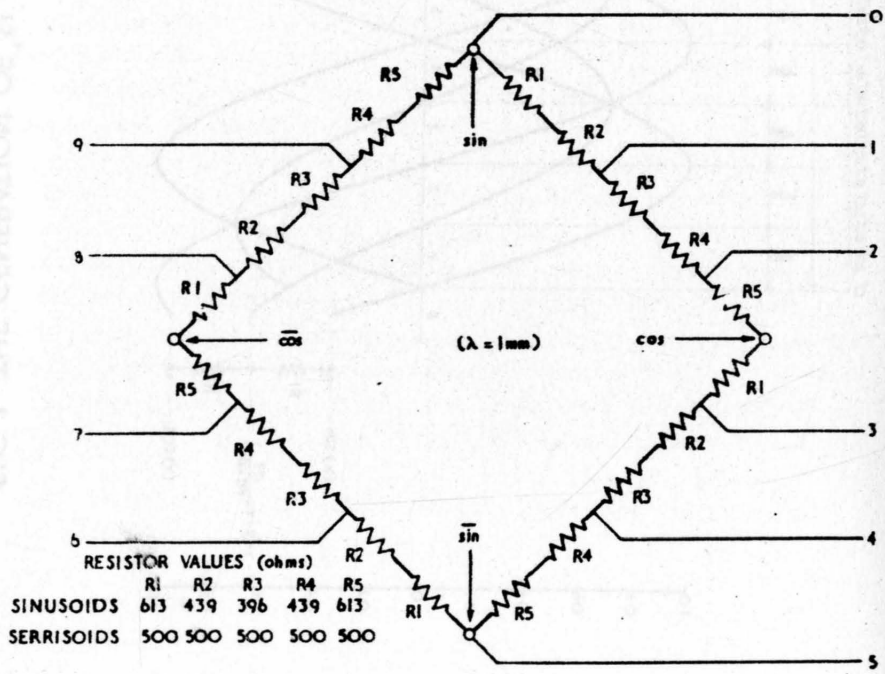


FIG 1 THE GENERATION OF 'SLAVE' WAVEFORMS



(a) 180° POTENTIOMETER



RESISTOR VALUES (ohms)

	R1	R2	R3	R4	R5
SINUSOIDS	613	439	396	439	613
SERRISOIDS	500	500	500	500	500

(b) 360° POTENTIOMETERS

FIG 2 sin/cos POTENTIOMETERS

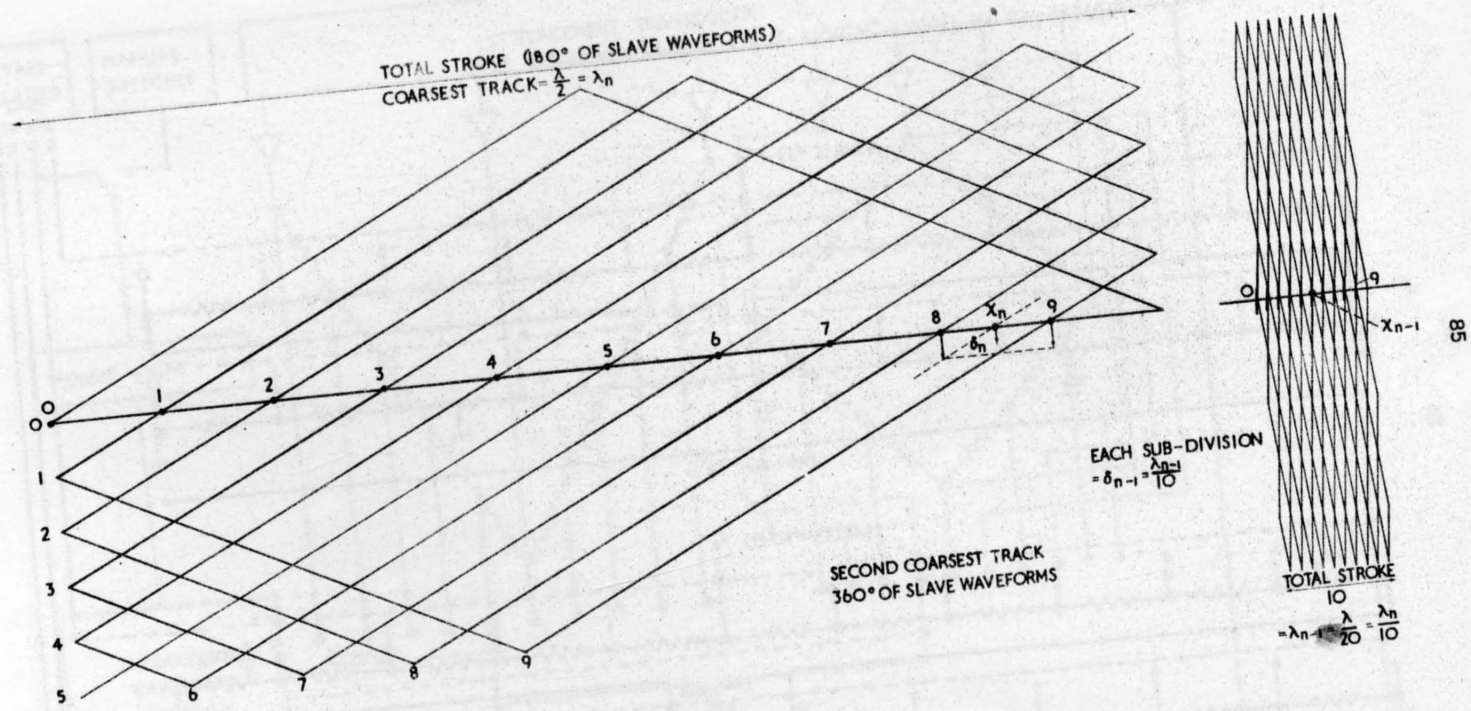


FIG 3 WAVEFORMS OF ÷ 10 SYSTEM

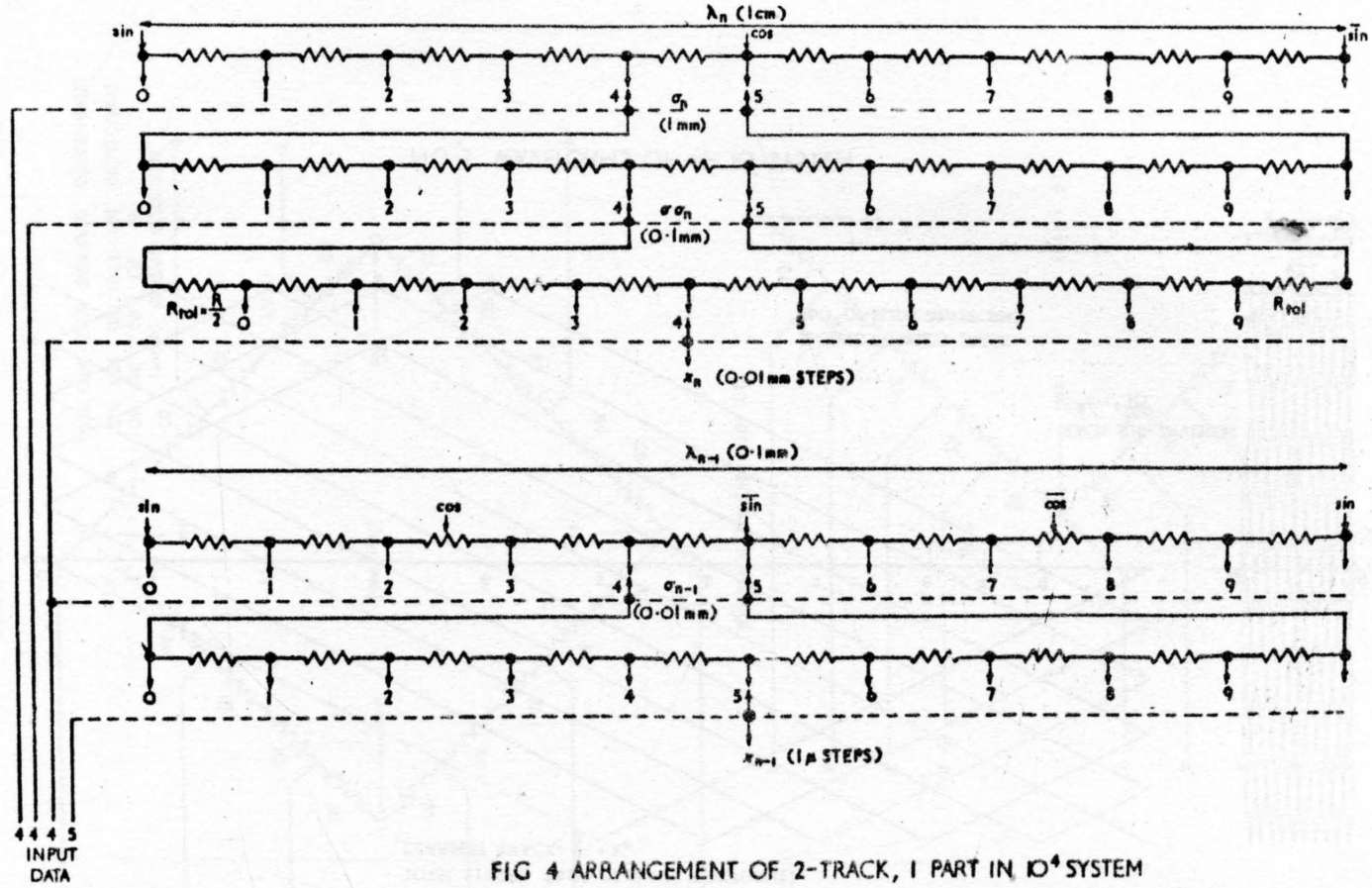


FIG 4 ARRANGEMENT OF 2-TRACK, 1 PART IN 10^4 SYSTEM

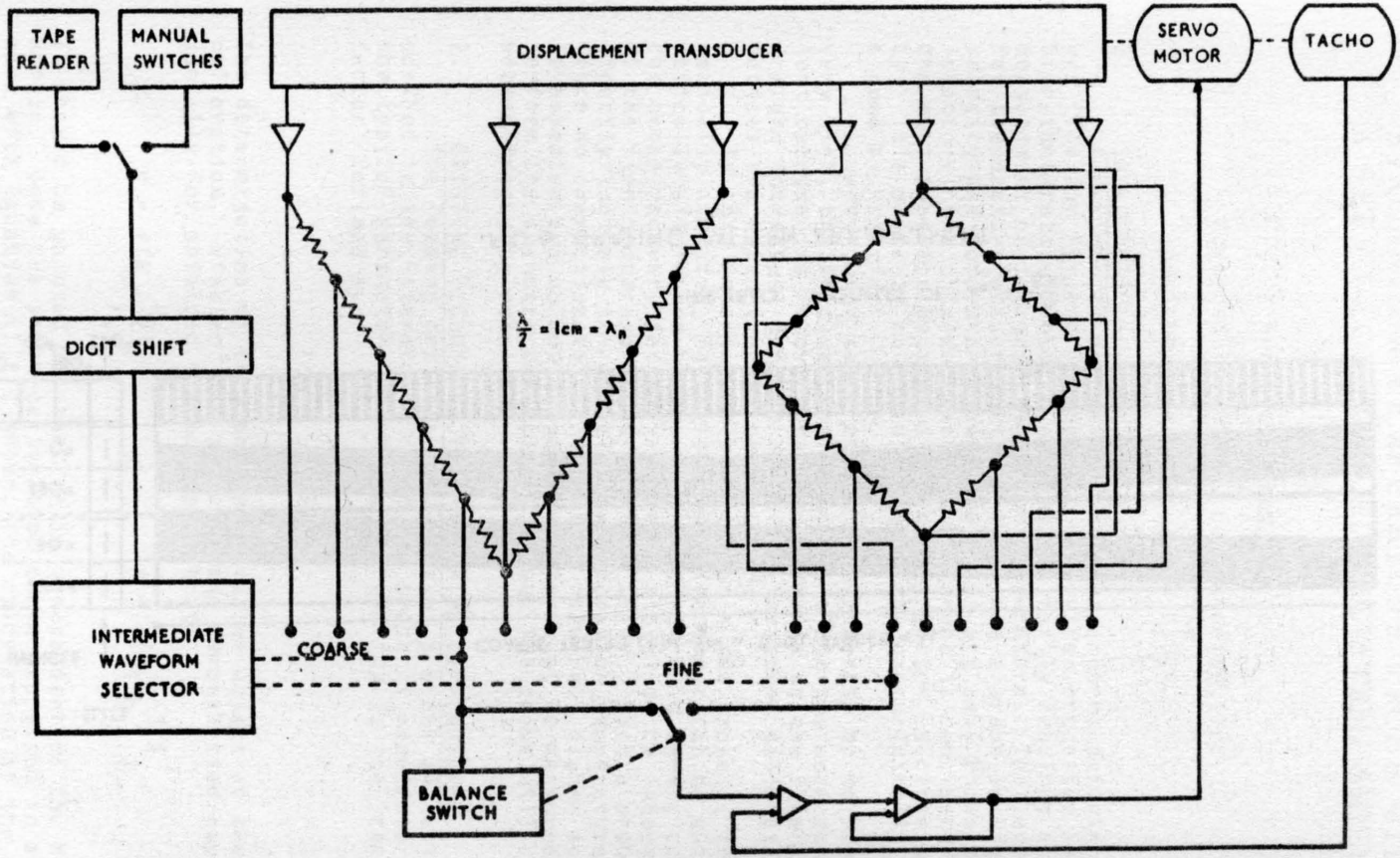


FIG 5 SCHEMATIC DIAGRAM OF COMPLETE CONTROL SYSTEM

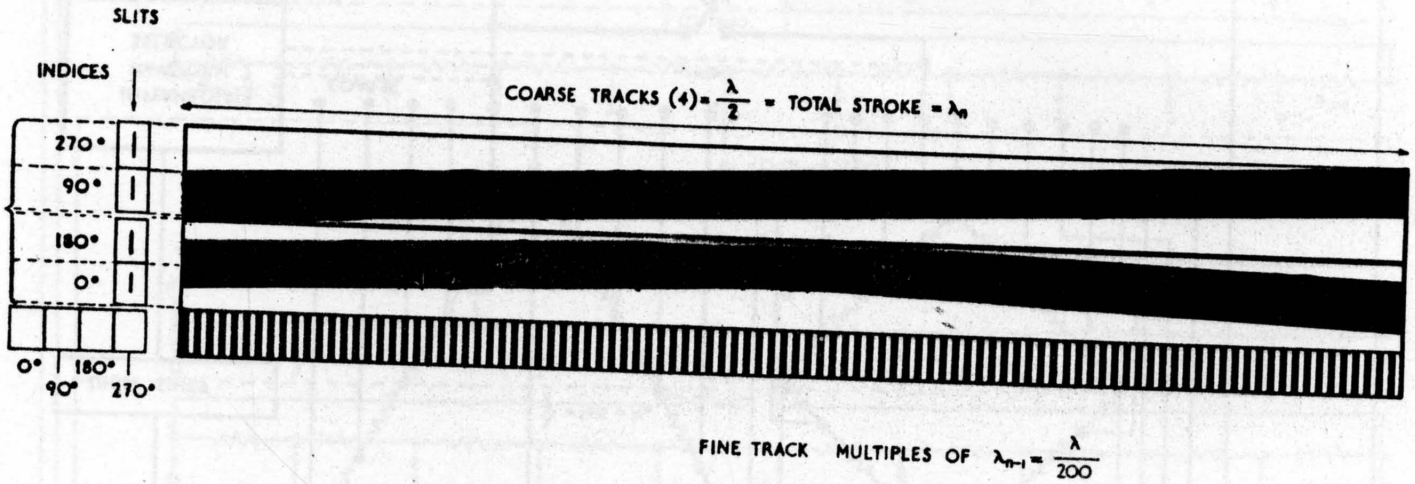


FIG 6 GRATING PATTERN ARRANGEMENT

STAGNATION TEMPERATURE PROBE WITH A SERVO-CONTROLLED RADIATION SHIELD

By

Horacio Andres Trucco

Advanced Technology Laboratories, Inc.
400 Jericho Turnpike
Jericho, New York, 11753 U.S.A.

1. INTRODUCTION

Experimental determination of parameters defining the gas flow for hypersonic regimes, combustion processes and high-temperature flows in general, primarily requires an exact measurement of the gas stagnation temperature. The techniques for temperature measurements by immersing a probe in high-velocity and moderate-temperature flows are well known and satisfactory; the shielded thermocouple probe is considered to be the most adequate because its recovery factor is close to unity. This probe is an evolution from the one originally designed by Franz[1].

However, for temperatures from moderate up to high or very high, the shielded thermocouple is not the most convenient because its recovery factor decreases as the flow temperature increases, and, in addition, the probe calibration devices are expensive and complicated.

In this report, the behavior of a new probe is theoretically analyzed, for which radiation error can be automatically eliminated by means of an electronic servo, causing the probe recovery factor (at steady state) to be independent of the flow temperature and equal to unity. Such a probe will directly indicate the actual flow stagnation temperature. Since no correction to raw output is required, higher accuracy is expected and no additional parameter must be measured to perform the standard corrections; also, the probe will show better frequency response.

2. EXISTING TECHNIQUES FOR MEASURING TEMPERATURES

The leading criteria in designing a probe are well described in the literature. Once the probe is manufactured, the typical parameter indicating its quality is the "recovery factor" defined by

$$r = \frac{T_s - T_\infty}{T_o - T_\infty}$$

The determination of this parameter is the object of the probe calibration. Winkler[2] has proposed and demonstrated experimentally for shielded thermocouples that:

$$r = f\left(N_u \frac{k_{gi}}{k_w}\right) \quad \text{or} \quad r = f\left(\rho_i T_o^{-3/4}\right) \qquad (1)$$

It should be pointed out that r is independent of the Mach number. Once the probe calibration curve expressed in accordance with Equation (1) is known, the calculation of the actual

stagnation temperature immediately comes from

$$T_o = \frac{T_s}{r + (1-r)\left(1 + \frac{\gamma-1}{2} M_\infty^2\right)^{-1}} \qquad (2)$$

The application of Equations (1) and (2) requires the simultaneous measurement of M_∞ and a pressure; these two additional measurements would not be required if r = 1.

To better point out the timportance of r, Figure 1 shows the ratio $\frac{T_s}{T_o} = f(M_\infty)$ plotted for different values of r. From these curves, we see how important a high r is, which enables us to make a better interpretation of the raw readings, and the T_o computation in the iteration process is simplified. Furthermore, the errors introduced by the uncertainty of γ decrease.

Everything which has been said above is valid for steady measurements; the unsteady phenomena add a new error in the measurement due to the inertia associated with the probe in following the temperature changes. Such errors depend upon the probe time lag associated with the time constant τ of the probe which is calculated as follows, if the radiation effect is negligible:

$$\tau = \frac{\rho_s c_s V_s}{h_s A_s}$$

If the radiation effect is to be taken into account, the time constant is expressed as follows:

$$\tau = \frac{\rho_s c_s \; D_{s/4}}{h_s + 4\sigma\varepsilon \; T_{s_f}^3}$$

This last equation is in accordance with a linearized solution[3], however, a different equation[4] can be used in calculating τ. The frequency for which the attenuation is 3 db is

$$f_c = \frac{1}{2\pi\tau}$$

For the smallest commercial thermocouples available and under the best flow conditions, f_c can only reach approximately 10 cycles per second.

It is always possible to reconstruct the actual temperature, even for unsteady behavior, if r and τ of the probe are known. Reference (6) gives an analogical technique to partially obtain it. However, when we wish to study fluctuations associated with turbulence or when obtaining correlation functions, we must consider

Probe's space resolution = Frontal cross section of the probe at the measurement point x Resolution length

where

$$\text{Resolution length} = \frac{\text{Mean flow velocity}}{2 \text{ Upper frequency limit of the probe}}$$

Generally, f_c is adopted as the upper frequency limit. Ideally, the resolution length of the probe should be approximately equal or less than its minimum frontal length in order to give a meaning when speaking of measurement at a point. A calculation under the best flow conditions for the smallest thermocouples gives a resolution length of the order of inches, which indicates that the thermocouple is inadequate for this type of investigation. At present, only the hot wire with an aspirating probe is adequate, but it is necessary to know simultaneously the gas stagnation pressure. It should be pointe out that the output of such a hot wire probe is influenced by fluctuations in the gas composition; therefore, the indication does not depend only upon the temperature. However, when this effect is important (for example, for combustion processes where oxidant, fuel and combustion products sensibly differ in their physical properties), the double hot wire aspirating probe can be used to separate temperature and composition fluctuations. The above review shows how useful and interesting it would be to have a thermocouple probe whose response would be in the range of the hot wire.

The radiation error in a shielded thermocouple probe may be calculated by

$$\text{error} = \frac{\sigma \varepsilon}{h_s} (T_r^4 - T_s^4)$$

or by a different formula suggested by Scandron et.al.[7]. Under steady conditions, the inside temperature of the shield T_r is always lower than the one reached by the thermocouple (except for surrounding enviroments whose temperature is higher than the gas stagnation temperature). The higher the gas flow temperature, the higher the difference will be, thereby corresponding to an increase in the radiation error. On the other hand, for transient phenomena--where the shield has a time constant higher by several orders of magnitude than the thermocouple--the difference between T_r and T_s will be still higher. The technique for reducing errors by adding several concentric shields[7,8] has given good results; however, this obviously is not the most desirable solution for experimental investigations due to the large size of the probe and the large time constant of the assembly.

3. DESCRIPTION OF THE PROPOSED SYSTEM

To the shielded thermocouple prove, we introduce the basic modification that the inside surface of the shield is automatically heated in such a way that its temperature T_r equals that of the sensible thermocouple T_s. This automatically eliminates the radiation error and in addition sensibly decreases conduction losses. As a result, the r of the probe will be independent of $T_{oe}(t)$ and should be very close to unity. This scheme from a steady point of view and using manual control has been already analyzed by Moffatt[9] and Wood[10]. Wood has demonstrated experimentally that with an adequate heating for the shield and thermocouple base it is possible to obtain a recovery factor equal to unity. However, the dynamic behavior for automatic control will be analyzed as follows: Figure 2 schematically represents the physical layout of the

proposed system. The thermocouple T_s indicates the flow stagnation temperature. The inside surface of the shield is coated with a heater element whose temperature may be varied by Joule's effect. This temperature is sensed by the thermocouple T_r. The signal of both thermocouples is amplified and its difference plus $\tau\, T_s(t)$ drives the current amplifier which feeds back to the heater.* The tachometer compensation network was added expressly to make the following analysis more general.

3.1 DISCUSSION OF EQUATIONS GOVERNING THE SYSTEM

We shall assume some simplifying hypotheses: First of all, we impose the design condition that between the radiation shield (or internal heater) and the outside surface of the shield, a thermal insulator such as quartz, pyrex, silica, etc., is located. Due to the fact that the thermal diffusion of these insulators is aboult 400 times less than most of the metals, we may assume with negligible error that: a) the heat transfer from the heater to the outside surface of the insulator is negligible** and therefore, the thermal balance of the insulator may be ignored, b) the temperature gradient in the radiation shield is zero in any direction. We assume a perfect thermal contact between the thermocouple T_r and the radiation shield, i.e. that the thermocouple will instantaneously respond to temperature changes in the radiation shield. In addition, we assume that the flow passing inside the probe is frozen but not adiabatic, that the catalyzer effect is absent, that the emissivity and absorptivity of the gas is negligible (this is an adequate assumption due to the small thickness of gas through which the radiation heat interchanges), and that the radiation shield completely surrounds the thermocouple. Finally, we assume that the heat conducted through the thermocouple wires is negligible. This assumption is expected to be adequate, because the support for the wires may be designed to allow the wire temperature to be close to that of the radiation shield and, at the same time, to take into account the well-known design criterion of length increasing and cross section decreasing of the exposed thermocouple wires. On the other hand, if we do not call for this assumption, partial derivatives should be introduced in the equation system. In order to simplify the understanding of the problem and the discussion of the equations, Figure 3 shows a block diagram from a functional viewpoint according to the previously established assumptions.

---

* It should be realized that this amplifier must be of the "clipper" type (so that it gives output for input>0). If not, the system would be unstable.

** This assumption is valid if the phenomenon analyzed has a time duration much smaller than the time constant of the thermal insulator.

Thermocouples T_s and T_r are assumed to be of the same material; however, due to the uncertainty concerning the equal behavior of both, there exists the possibility for T_r to have an output lower than T_s. It is obvious that for steady regimes this could introduce a system instability. In order to eliminate this possibility, it is necessary to introduce a constant τ which is of the order of the mentioned uncertainty. For the same purpose, we have intentionally introduced independent amplifiers for T_s and T_r whose gain could differ by the order of hundredths ($K_r > K_s$). Obviously, when τ is higher or the difference between K_r and K_s is higher, the radiation error introduced will be larger. Under these assumptions, the equations governing the system are as follows:

Heat balance of the thermocouple:

$$a \frac{dT_s(t)}{dt} = b \{T_{o_i}(t) - T_s(t)\} + c \{T_r(t)^4 - T_s(t)^4\} \quad (3)$$

Heat balance of the radiation shield:

$$d \frac{dT_r(t)}{dt} = e \left\{\frac{T_{o_e}(t)+T_{o_i}(t)}{2} - T_r(t)\right\} - r\, 1(t)^2 - c \{T_r(t)^4 - T_s(t)^4\} \quad (4)$$

Energy balance of the flow circulating inside the probe (we assume there is no variation of kinetic energy).

$$T_{o_i}(t) = T_{o_e}(t) = f\left\{\frac{T_{o_e}(t)+T_{o_i}(t)}{2} - T_r(t)\right\} \quad (5)$$

Electronic feedback system:

$$I(t) = K \{T_s(t) - K_1 T_r(t) - T\} - \tau \frac{dT_s(t)}{dt} \quad (6)$$

where

$$K = \frac{K_f K_r q}{R} \qquad \text{gain} \quad (7)$$

$$K_1 = \frac{K_r}{K_s} \qquad \text{gain ratio} \quad (8)$$

$$\tau = \frac{K_f \tau_1 q}{R} \qquad \text{time constant} \quad (9)$$

and T= constant, obtainable by adjusting the reference junction of the thermocouple T_r. The significance of the remaining coefficients appears in the next section.

Since the resulting system is non-linear, we cannot analyze its behavior by means of standard servomechanism techniques; a linearization of the above set of equations has been used allowing standard servo techniques to be applied. Since the results that could be obtained by this method were limited, it was deemed necessary to solve the problem in an

analog computer.

4. SYSTEM RESPONSE FROM ANALOG COMPUTATION

The set of Equations (3) to (6) governing the proposed system was solved by analog computation. The investigated solutions correspond to a typical probe geometry used for experimental research (see Figure 4). For this probe the equation coefficients were computed as shown below. For an optimum value of the shield vent holes, we assume a Mach number inside the probe, $M_i = .10$; for the flow stagnation temperature, we adopt T_{oe}(average)=2,000°F. The stagnation pressure inside the probe is equal to the pitot pressure since the common values in hypersonic regimes vary between 1 and 50 psia; for our case, we adopt 15 psia. The coefficients are not to be assumed constant, because they depend upon free flow conditions, however, we accept that they are constant within a limited range around the average values. For the probe geometry, probe materials and flow conditions adopted, the following values have been determined for the coefficients:

$$a = \rho_s c_s V_s = 1.50 \times 10^{-10} \quad \left(\frac{BTU}{°f}\right)$$

$$b = h_s A_s = 1.50 \times 10^{-8} \quad \left(\frac{BTU}{sec\,°F}\right)$$

$$c = \sigma\varepsilon\alpha A_s = 3.00 \times 10^{-20} \quad \left(\frac{BTU}{sec\,°F}\right)$$

$$d = \rho_r c_r V_r = 1.50 \times 10^{-6} \quad \left(\frac{BTU}{°F}\right)$$

$$e = h_r A_r = 8.50 \times 10^{-6} \quad \left(\frac{BTU}{sec\,°F}\right)$$

$$f = \frac{h_r A_r}{w\, c_p} = 0.90$$

$$r = RE = 3.16 \quad \left(\frac{BTU}{sec\ Amp^2}\right)$$

$$R = 3{,}330\ (\Omega)$$

The computation of a,c,d, and P is obvious, however, the remaining computations require more attention. In our case, b has been computed by adopting the values of h_s for heat transfer between air and spheres according to Figures 10-11, page 266, of McAdams[5]. For calculation of e and f, a laminar flow has been assumed inside the probe (since the Reynolds number referred to the insdie diameter is ≃ 450) adopting for h_r the experimental values of Figures 9-20 from McAdams[5]. When employing the thin film thermometer technique to manufacture the radiation shield with conducting paints, its thickness could be reduced up to 100 times less than the typical value adopted in Figure 4. Consequently, a value of

$d=1.50\times10^{-8}$ was also adopted for the computations. Two basic tests were investigated: step function and periodic variation of the flow stagnation temperature assuming an initial temperature of 2,000°F.

4.1 STEP FUNCTION RESPONSE

The basic step function test investigated was $T_{0e}(t)=460\ u(t) + 2000^{o}F$; however, two steps with amplitude 615°F and 50°F were also investigated in order to find the effect of the input amplitude on the output response. The response $T_s(t)$ of the standard shielded thermocouple probe, that is, without the compensating electronic servo proposed, (or for $K=o$, $\tau=o$) shows a settling time* of 1.00 second:** A series of tests with the computer showed that the settling time can be reduced substantially by reducing the radiation shield time constant, consequently, all results presented here correspond to a thin radiation shield ($d=1.50\times10^{-8}$).

Figure 5 shows the probe response (with thin radiation shield) for the four cases investigated: standard probe, proposed probe with compensated output, etc. This figure clearly shows the improvement in the standard thermocouple settling time, that is, from 1.0 seconds it was reduced to 0.003 seconds, a reduction of 333 times. Curve D was obtained by restoring the output by means of a derivative compensating network as shown by the dotted lines in Figure 3, by proper choice of the time constant τ_c.

The effect of K_1 (see Equation (8)) was also investigated; it was found that by making $k_1>1$ the settling time increases, but no steady state error was observed.

4.2 RESPONSE TO PERIODIC INPUT

The temperature in a gas flow does not stay steady, but always shows random, or sometimes periodic, fluctuations. In order to investigate this effect on the probe response, two periodic imputs (superimposed to the initial 2000°F) were chosen: sinusoidal waves and square waves.

Figure 6 shows the probe response for five sinusoidal inputs of the same amplitude: 460°F, and 50, 100, 200 and 800 cycles per second. The main observation from Figure 6 is the presence of steady state error due to the sinusoidal input. This somewhat unexpected result is due to the presence of the clipper amplifier on the feedback loop, which causes the cooling cycle of the radiation shield to be much slower than the heating cycle.***

---

*"Settling Time" is the time required to reach 99% of the steady state output.

**This large settling time is typical for rugged probes which were designed for hours of satisfactory operation.

***If thermoelectric or "Peltier" cooling effect could be used on the radiation shield for the cooling cycle, the probe would not show steady state error.

This behavior is undesirable; the steady state error is slightly dependent upon input frequency, and it averages 39% of the periodic amplitude. Figure 6 also shows the steady state attenuation and the fact that the response is periodic with the same frequency; Lissajou's path obtained during the investigations showed this conclusion more clearly. Square wave inputs give the same type of steady state error as for sinusoidal inputs.

5. CONCLUSIONS

The standard shielded thermocouple probe was modified by adding an electric heater on the inside surface of its shield; this new type of radiation shield is fed and controlled by an electronic servo as described in Section 3. After deriving the system governing equations and analyzing its response with an analog computer, we conclude that the proposed probe system is characterized by the following;

a) The probe system shows no radiation error after reaching steady state under any type of input. Then its recovery factor should be unity.

b) If only signle loop feedback is used (no tachmeter feedback included) to control the radiation shield heater, the probe system is stable under any type of input independently of the gain K.

c) When using the complete feedback system, the probe system is stable for any type of input provided the gain K and the tachometer time constant τ stay below certain stability limits, which vary with input amplitude.

d) The settling time of the standard shielded thermocouple (no servo) is 1.0 second for a step amplitude of 460^{o}F over the initial air flow at 2000^{o}F.

e) The settling time of the proposed probe system for positive step function is shorter when the radiation shield heater is thinner.

f) By using thin film radiation shield (0.00001 inch thick) and adding a derivative compensation network on the output, the probe system settling time (for a positive step) was reduced to 0.003 seconds, which means a 333 times faster response than the standard shielded thermocouple.

g) The probe system response to a "negative step" (referred to the initial 2000^{o}F) is the same as the standard shielded thermocouple; this is due to presence of the clipper amplifier on the feedback loop. If the state of technology would permit use of the Peltier cooling effect in conjunction with the radiation shield, the probe response would be the same for positive and negative steps; furthermore, the

two undesirable conclusions h) and I) would not exist.

h) Under periodic input (superimposed to the initial 2000°F), the probe system response shows a steady state error of 39% of the input amplitude. The time required to build up this error is 25 milliseconds. This steady state error is not very important because for a 2000°F gas flow, typical temperature fluctuations average 50°F of amplitude. This will give 20°F steady state error, which is only 1% of the reading.

i) The probe system attenuation to an 800 cycles per second sinusoidal input is 95%. This fact shows that the probe is not adequate to measure temperature fluctuations associated with turbulence, etc.

11

6. RECOMMENDATIONS

The investigated probe geometry is quite rugged, and can be used in standard supersonic and hypersonic facilities. Because its response was found to be 3 milliseconds this probe should be fabricated and then its actual behavior compared with analytical results. The scope of the present investigation should be extended further; i.e., since the probe response depends basically upon the time constants of the thermocouple and radiation shield, an investigation should be conducted on the use of a "thin film thermocouple" as well as a vacuum deposited radiation shield in order to improve the probe response and minimize the steady state error due to fluctuating inputs. The last approach may lead to the design of a stagnation temperature probe whose response could match the hot wire response.

SYMBOLS

A	Area, wetted by the flow
D	Internal diameter of the probe, or according to subscript
E	Conversion factor from Watts to BTU/sec.
M	Mach number
Nu	Nusselt number=$h_s D_s / k_{gi}$ (for the sensible thermocouple)
R	Electrical resistance of the radiation shield
T	Absolute Temperature
V	Volume
c	Specific heat
c_p	Specific heat at constant pressure
h	Convection heat transfer coefficient
k_g	Thermal conductivity of the gas
k_w	Thermal conductivity of the thermocouple wires
q	Thermoelectric effect
r	Recovery factor of the probe
t	Time
w	Mass rate of flow inside the probe
α	Absorptivity
γ	Ratio of specific heats
ε	Emissivity
σ	Stefan-Boltzmann constant
ρ	Density
τ	Time constant

$u(t)$ Unit-step function(defined for $t>0$)

SUBSCRIPTS

o denotes adiabatic stagnation values of the free stream

∞ denotes free stream conditions

o_e denotes adiabatic stagnation values of the flow at the entrance of the probe

o_i denotes real stagnation values of the flow at the closed end of the probe (around the sensible thermocouple)

s denotes properties or geometry of the sensible thermocouple

r denotes properties or geometry of the radiation shield

i denotes properties of the gas at the closed end of the probe (around the sensible thermocouple)

REFERENCES

1. Franz, A. - Pressure and Temperature Measurement in Supercharger Investigation. NACA TM953, September 1940.
2. Winkler, E M. - Stagnation Temperature Probes for Use at High Supersonic Speeds and Elevated Temperatures. NAVORD Report 3834, October 1954.
3. Baldwin Lima Hamilton - HT Micro-Miniature Thermocouple, Technical Data 4336-1.
4. Scandron, M.D., Warshawsky, I. - Experimental Determination of Time Constant and Nusselt Numbers for Bare Wire Thermocouples in High Velocity Air Streams and Analytical Approximation of Conduction and Radiation Errors. NACA TN2599, 1952.
5. McAdams, W.H. - Heat Transmission. McGraw-Hill, 3rd. Edition, 1954.
6. Grote, H.H. - Improvement of Sensor Response by an Analog Technique. USAELRDL Technical Report 2445, March 1964.
7. Moffatt, E.M. - Multiple Shielded High Temperature Probe-Comparison of Experimental and Calculated Errors. SAE Preprint No. T 13, January 1952. Also in SAE Quart. Trans. Vol. 6, 1952
8. Kudryavtsen, E.V. - Thermocouple Measurement of the Temperature in a High-Velocity Gas Flow,Gas Dynamics and Physics of Combustion-Translated from Russian, Israel Program for Scientific Translations, Jerusalem 1962.
9. Moffatt, E.M. - Errors in High Temperature Probe for Gases. ASME Paper 48-A-52, December 1948.
10. Wood, R.D. - A Heated Hypersonic Stagnation Temperature Probe. Jour. Aero/Space Sciences, Vol. 27, July 1960.

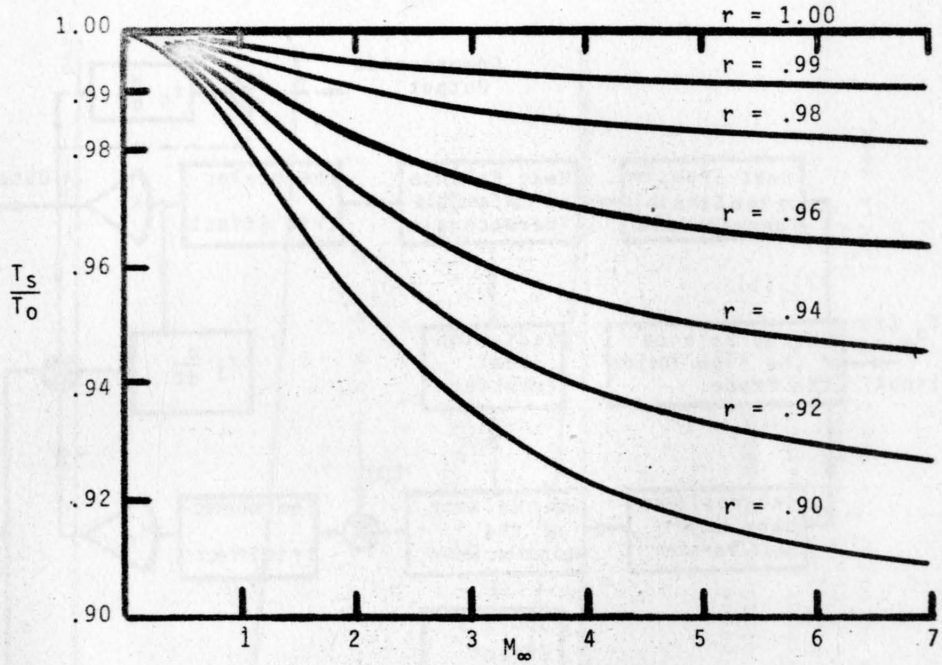


Figure 1: Effect of Probe Recovery on Actual Output (γ=1.40)

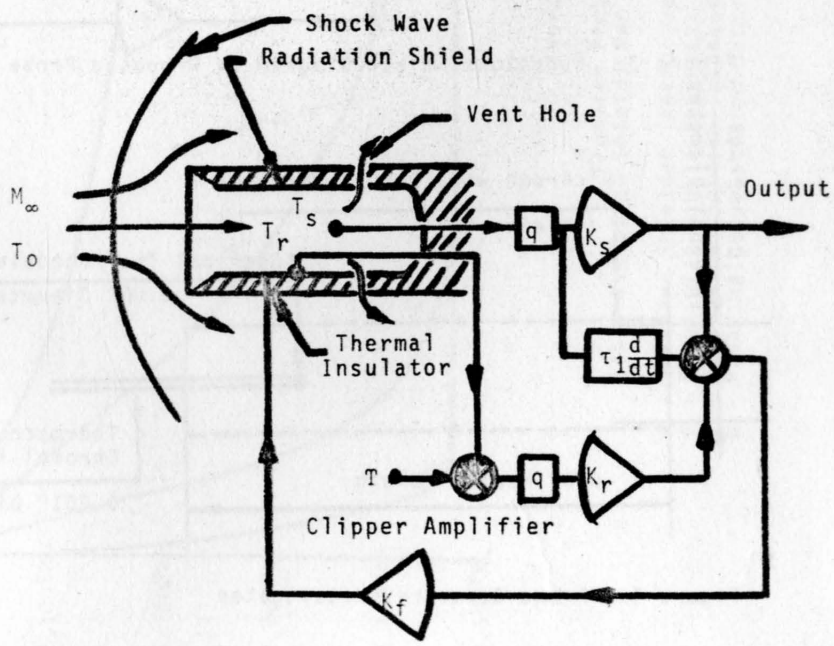


Figure 2: Schematic of Proposed Probe With Electronic Servo

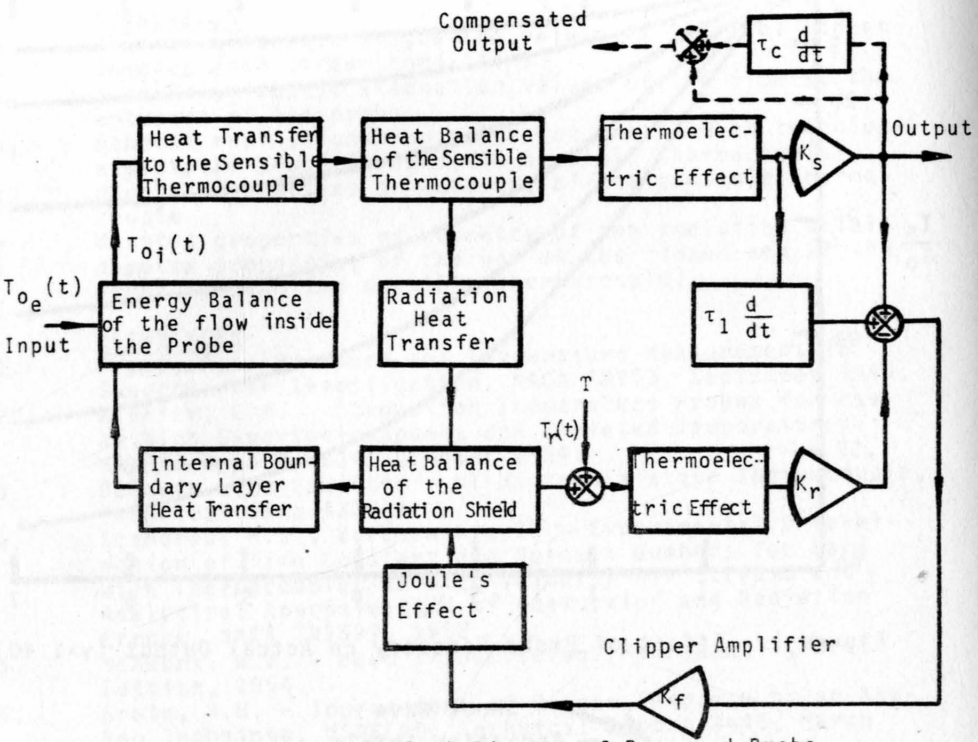


Figure 3: Functional Block Diagram of Proposed Probe

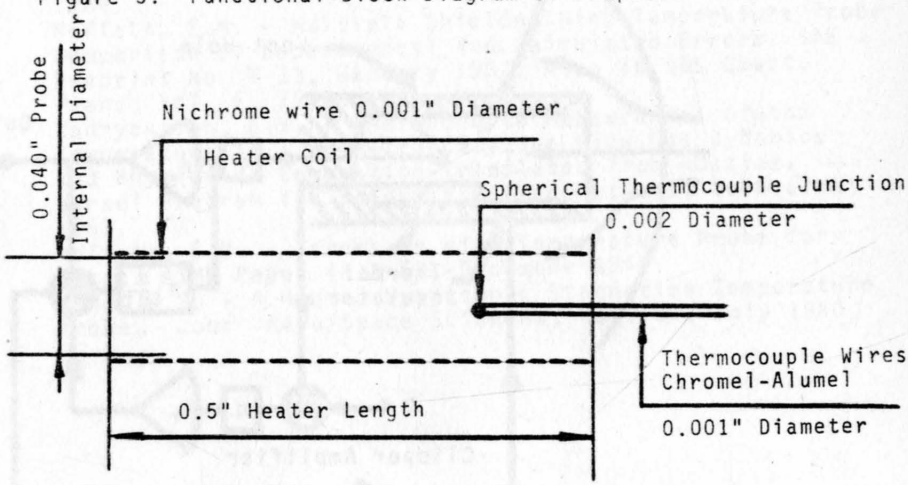


Figure 4: Probe Geometry Investigated

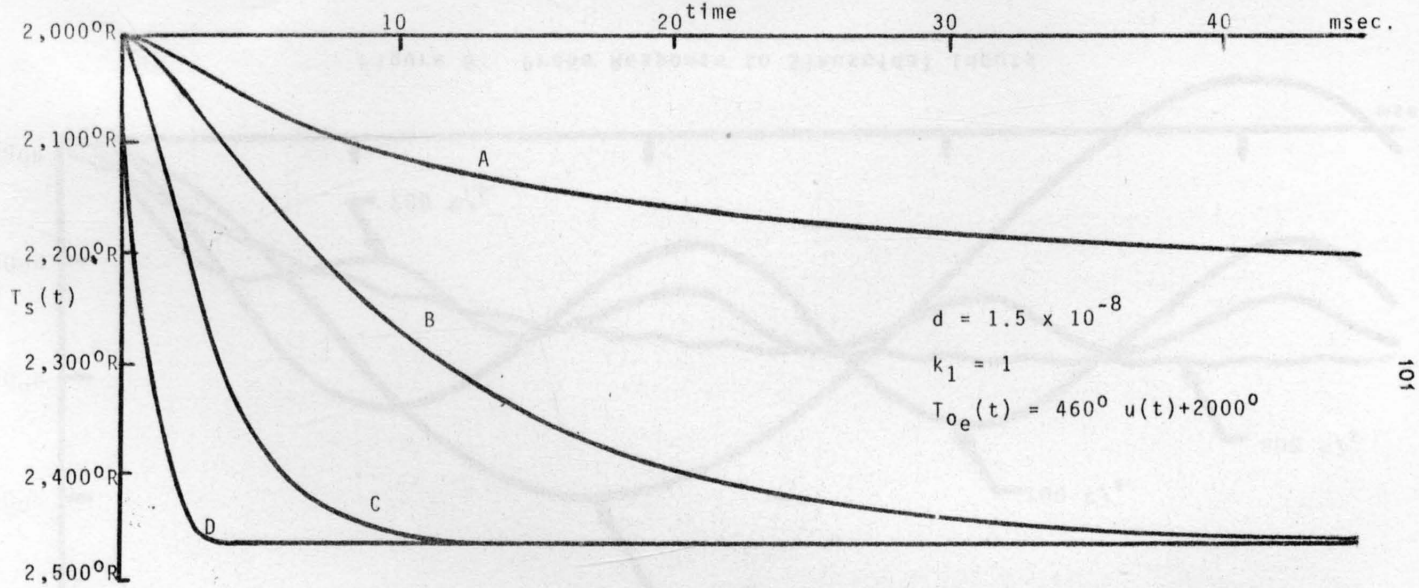


Figure 5.

Proposed Probe Response
(Thin Radiation Shield) for four cases investigated:

A - Standard Shielded Thermocouple Without Compensation at all.
B - Probe with Simple Feedback: K = 6.85×10^{-4} - τ=o - τ=o
C - Probe with Complete Feedback: K = 6.85×10^{-4} - $\tau = 8.14 \times 10^{-6}$ - τ_c=o
D - Probe with Compensated Output: K = 6.85×10^{-4} - $\tau = 8.14 \times 10^{-6}$ - τ_c=2.4 msec.

$T_{o_e}(t) = \{460^o \sin(2\pi ft)\}\ u(t)+2000^o$

$d = 1.5 \times 10^{-8}$

$K_1 = 1$

$K = 6.85 \times 10^{-4}$

$\tau = 8.14 \times 10^{-6}$

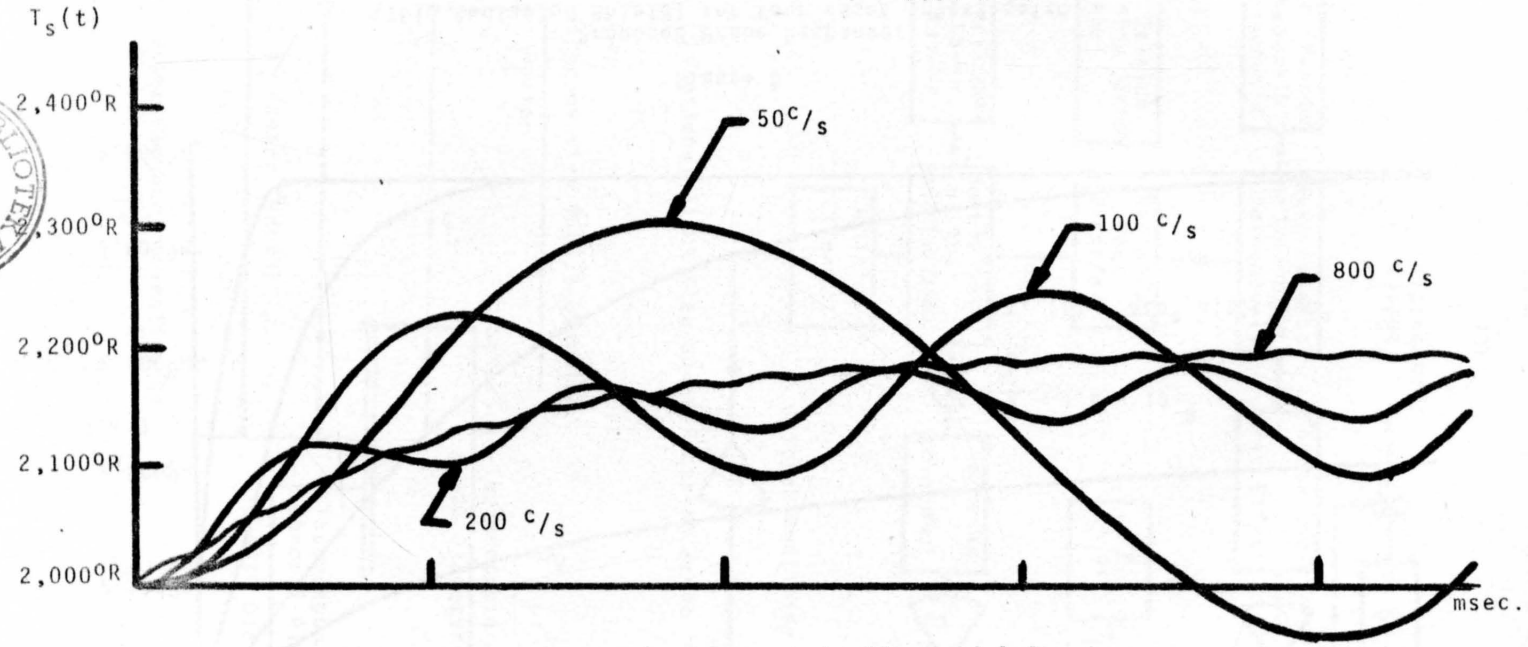


Figure 6: Probe Response to Sinusoidal Inputs