IFAC

INTERNATIONAL FEDERATION
OF AUTOMATIC CONTROL

WARSZAWA 1969

Theoretical Aspects of Devices

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

TECHNICAL SESSION
64

NOT

Organized by
Naczelna Organizacja Techniczna w Polsce

INTERNATIONAL FEDERATION OF AUTOMATIC CONTROL

Theoretical Aspects of Devices

TECHNICAL SESSION No 64

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. 79/69.

ОБ ОДНОЙ СИСТЕМЕ АВТОМАТИЧЕСКОГО УПРАВЛЕНИЯ ПРОЦЕССАМИ ИЗГОТОВЛЕНИЯ МИКРОСХЕМ

Институт кибернетики АН УССР В.М. Глушков, В.П. Деркач,
Киев, СССР Г.Т. Макаров

За последние 10-15 лет количество электронных вычислительных машин, являющихся одним из наиболее мощных средств повышения эффективности разнообразной деятельности человека, возросло в сотни раз. В то же время они как правило собираются вручную, что повышает их стоимость, снижает надежность, ограничивает применение. Время, требуемое на разработку и изготовление машин, бывает соизмеримо, а иногда даже больше времени морального их старения, которое постепенно сокращается и в настоящее время составляет приблизительно 3-4 года.

Несмотря на многочисленные попытки исследователей автоматизировать процесс производства схем, построенных на навесных деталях, из-за быстрого усложнения аппаратуры и большого разнообразия используемых деталей значимых результатов получено не было. Потребовалось отыскание существенно новых физико-технологических путей построения этой аппаратуры, которое привело к замене отдельных деталей и компонентов, используемых в качестве строительных элементов, известными теперь интегральными схемами, выпускаемыми в виде тонких пленок или твердых полупроводниковых монолитных блоков.

Наряду с резким уменьшением габаритов и повышением долговечности приборов важнейшим достоинством таких схем является упрощение задачи автоматизации этапов их производства.

Хотя в основу микроэлектроники положены достаточно хорошо изученные физические явления и такие известные методы, как вакуумное осаждение вещества на подложку, диффузия, рекристаллизация, окисление, эпитаксиальное выращивание монокристаллических пленок, травление, фотолитография, термообработка и др., практическое использование их усложняется из-за необходимости пространственной локализации происходящих процессов на многочисленных микронных участках материала. Разработка способов осуществления строго контролируемого селективного протекания физико-технологических процессов, поэтому, является важнейшей проблемой этой новой науки.

Одним из таких методов, который сейчас чаще всего применяется на практике, является воздействие на вещество через

маску /трафарет/. Использование его привело к значительным результатам, окончательно утвердившим микроэлектронику как прогрессивную науку.

Но при использовании этого способа выявились и трудности. В камеру приходится помещать трафареты различных конфигураций и механизмы их перемещения, которые нередко требуется нагревать для обезгаживания и получения глубокого вакуума. При многократном совмещении трафаретов возникают ощутимые ошибки и, следовательно, снижается точность геометрических размеров изготовляемых компонентов, ухудшается воспроизводимость их характеристик. Во многих случаях бывает вообще невозможно изготовить трафарет нужной конфигурации.

Точность размеров повышает за счет использования прецизионных контактных масок, изготовляемых на обрабатываемых подложках методами фотолитографии. Но использование фотолитографических процессов при обработке полупроводниковых материалов приводит к увеличению разнотипности применяемых технологических операций, что влечет за собой появление дополнительных ошибок. Из-за случайного распределения бракованных компонентов по пластине рисунок монтажа в каждом экземпляре интегральной схемы, сделанный таким образом, получается индивидуальным. Следовательно, требуется изготовление новой маски для каждой пластины. Чтобы избежать резкого удорожания продукции и больших потерь времени на это, сейчас ограничивают число элементов, изготовляемых на одной пластине. И все равно выход годных изделий в лучшем случае не превышает нескольких десятков процентов.

Большое разнообразие операций, значительная часть из которых, к тому же, выполняется вручную, существенно затрудняет решение задачи полной автоматизации производства микросхем. Следует учесть также, что быстро появляющиеся новые технологические приемы вызывают значительные переделки дорогостоящих автоматических линий. Поэтому во многих странах ведется усиленный поиск таких технологических приемов и инструментов, параметры и характеристики которых более полно соответствовали бы задачам создания устройств с элементами микронных размеров.

Обнадеживающие перспективы для успешного решения этих задач открывают результаты изучения вопросов взаимодействия электронных и ионных пучков с твердым телом. Электронная и ионная /элионная/ обработка материалов по праву считается сейчас наиболее совершенной из всех существующих методов создания компо-

нентов микросхем. С помощью электронных и ионных пучков сейчас изготовляются р-п переходы, транзисторы, производится микросварка, полимеризация, разложение химических соединений с целью восстановления химических элементов на локальных участках подложки, рекристаллизация, экспозиция фоторезистивных слоев, напыление пленок, скрайбирование, герметизация, измерение и контроль параметров технологических процессов и изделий, определение химического состава материалов и т.д. – выполняется множество основных операций, необходимых и применяемых фактически для построения интегральных схем.

Преимуществом элионных методов является возможность выполнения всех технологических этапов в вакуумной камере без ее разгерметизации, т.е. в условиях, обеспечивающих высокую воспроизводимость параметров изделий. Плотность мощности электронных и ионных пучков легко регулируется и достигает миллионов киловатт на $см^2$. С их помощью можно локализовать физико-технологические процессы в очень малом пространстве. Полученный практически минимальный диаметр, составляет доли микрона.

Поскольку процессы локализуются не путем механических перемещений элементов, скажем, трафаретов, а с помощью воздействующих на заряженные частицы электрических и магнитных полей, то для обработки можно почти мгновенно и с большой точностью выбирать любую точку изготовляемой микросхемы, получать какие угодно конфигурации обрабатываемых участков, отказавшись от масок или, если это надо, с наибольшей возможной сейчас точностью изготовлять сами маски.

Элионная технология представляет собой один из примеров такой области техники, быстрое развитие которой обусловлено появлением кибернетических средств управления. Вручную, с помощью оптических приборов можно создавать лишь лабораторные образцы элементов в единичных экземплярах. Невозможно производить в течение разумного времени таким способом целые блоки – многокомпонентные схемы с приемлемой воспроизводимостью их параметров.

С точки зрения разнообразия характера технологических процессов и количества элементов, на которые должна воздействовать управляющая система, элионные установки являются выгодным , мало изменяющимся во времени объектом управления, поэтому принципы автоматического управления и применяемые для их воплощения технические средства, если они выбраны с учетом новейших научно-

технических достижений, должны быть достаточно жизнеспособными.

В Институте кибернетики АН УССР на первом этапе разработана разомкнутая система автоматического управления процессами изготовления компонентов элионных интегральных схем /"Київ-67"/, которая сейчас эксплуатируется на одном из предприятий и уже показала высокую надежность и эффективность. В основу ее положен цифровой способ управления, как более экономичный по времени и обладающий лучшими возможностями, чем метод использующий непрерывную развертку.

При электронно-лучевом изготовлении компонентов микросхем, например, чаще всего бывает выгодным применение импульсного режима воздействия пучка на материал. Требуется задание длительностей импульсов и пауз, а также количество импульсов обработки в каждой точке. В случае непрерывной развертки всего растра, как это делается, скажем, при использовании фотокопира, длительности пауз всегда связаны с длительностями импульсов уравнением

$$t_п = \frac{LHt_и}{\mathcal{D}d_п}$$, где

L и H - длина и высота растра;

$t_и$ - длительность импульса обработки;

$\mathcal{D}$ - допустимое расстояние, на которое может переместиться пучок за время действия импульса обработки, получаемое из условий обеспечения достаточной разрешающей способности.

$d_п$ - диаметр пучка.

Как видно из этой формулы соотношение между $t_п$ и $t_и$ не может быть выбрано произвольным. Кроме того, облучаемые и свободные от воздействия пучка участки кадра развертываются с одной и той же скоростью. Вследствие этого приходится непроизводительно тратить время на движение "запертого" луча, которое накапливается с увеличением количества импульсов и разнообразия временных условий обработки и определяется из уравнения.

$$t_{з.л.} = \sum_{k=1}^{m} \sum_{l=1}^{n} a_{ik} \cdot t_{пk}$$, где

a - относительное число необрабатываемых точек на подложке;

m - число различных длительностей импульсов обработки, требуемых при изготовлении схемы;

n - наибольшее число импульсов одной и той же длительности, необходимое для обработки в одной точке растра.

Метод непрерывной развертки представляется приемлемым только тогда, когда нет необходимости в изменении временных режимов обработки в пределах растра. В подавляющем большинстве других случаев должна быть обеспечена возможность быстрой автоматической установки пучка в любой участок подложки и облучения выбранной точки неподвижным лучом при каком угодно отношении t_n/t_u .

Для осуществления этого наиболее удобна контурная развертка, получаемая преобразованием цифровых кодов, соответствующих координатам точек, в отклоняющие токи или напряжения. Использование шаговой контурной развертки приводит к необходимости применения для управления электроннолучевой установкой цифровой системы.

Имеющиеся универсальные цифровые машины не приспособлены для данной цели, поскольку они не в состоянии обеспечить минимальное, необходимое для реализации преимуществ электроннолучевой технологии быстродействие и требуют чрезмерного усложенения процессов программирования.

Нашу систему можно применять при лабораторных исследованиях и мелкосерийном производстве. Она может быть также использована для электроннолучевой сварки, фрезеровки и в других разнообразных случаях, когда нужно с большим быстродействием воздействовать на объект управления одновременно по нескольких каналам.

Одной из наиболее важных ее особенностей является простота программирования технологических задач. По одной команде может отрабатываться любая из пяти наиболее часто встречающихся геометрических фигур произвольных в пределах растра размеров /рис. I/. При этом энергетические и временные режимы облучения одинаковы для всех точек. Изменяя параметры α и β "точечного растра", используемого при изготовлении схем с регулярно расположенными компонентами /например, диодных матриц/, устанавливаются расстояния между обрабатываемыми точками. При $\alpha = 1$ или $\beta = 1$ получается серия параллельных линий, нужная, скажем, для создания так называемой токопроводящей змейки. Если $\alpha = 1$ и $\beta = 1$, то "точечный растр" преобразуется в прямоугольную площадку.

Для того, чтобы на одной пластине кремния можно было изготовлять большое количество регулярно расположенных схем по упрощенной программе, предусмотрена возможность обработки по одной команде серии регулярно расположенных площадок. Для этой цели служит кадр "ряд прямоугольников". Если этот кадр задать в пре-

делах одной строки, то на подложке воспроизводится прерывистая линия, встречающаяся при изготовлении резисторов, выполняемых в виде тонкопленочной змейки.

Отработка линий и площадей произвольной формы, требующихся, например, при сварке корпусов, изготовлении масок, засветке фоторезистивных слоев и т.п. осуществляется с помощью кадров "наклонная линия", "окружность" /"дуга"/ и "площадь". В последнем две стороны фигуры или одна из них может ограничиваться наклонной линией либо частью окружности, образуя треугольник, равностороннюю или прямоугольную трапецию, круг, сегмент и т.п.

Комбинируя эти геометрические фигуры и связи между ними на подложке, и задавая их отработку в нужной временной последовательности, можно создавать сложные схемы, строить разнообразные функциональные устройства.

Общий вид "Киева-67" дан на рис. 2, а блок-схема машины и функциональное устройство блоков БУО и БО - на рис. 3 и 4.

Каждая команда /код кадра/ представляется десятью двенадцатиразрядными двоичными словами, для хранения которых применено магнитное запоминающее устройство, где они расположены в постоянной последовательности. В этих словах содержатся данные об энергетических режимах пучка, временных параметрах обработки, а также указания о законе перемещения луча по поверхности площади и все исходные величины, необходимые для получения фигур требуемых размеров. Для задания начала и конца движения луча служат регистры Сч НХ, Сч КХ, РНУ и РКУ. При шаге 5 мк и дополнительном разряде реверса максимальная площадь обрабатываемой поверхности получается равной 10х10 мм.

Объем МЗУ 4096 двенадцатиразрядных слов, что составляет более чем 400 кодов кадров. Этого количества, повидимому достаточно для многих применений, тем не менее предусмотрена возможность оперативного пополнения или замены программ, содержащихся в памяти, в ходе выполнения технологических операций с помощью устройства ввода с перфоленты или другой вычислительной машины, используемой, например, для автоматического составления программ изготовления микросхем или их корректировки по сигналам обратных связей.

Характер и количество основных геометрических фигур оказали решающее влияние на построение вычислительного узла /БУО/, основу которого составляет перестраиваемый на отработку различных

кадров /в соответствии с кодом в регистре признака кадра/ линейно-круговой интерполятор, собранный на двух цифровых интеграторах. Каждый из них в свою очередь состоит из счетчика и накапливающего сумматора.

При отработке точечного растра, например, выраженное в количестве шагов расстояние a между точками по оси X задается кодом в СчX$_0$, а расстояние b по оси У - кодом в СчУ$_0$. Эти коды в течение расчета всех точек кадра хранятся в сумматорах ΣX и ΣY .

После облучения первой точки, время обработки которой определяется блоком формирования временных параметров /БФВП/, осуществляется переход к следующей, для чего изменяется на единицу код в счетчике отклонения СчX /блок отклонения, БО/ и одновременно вычитается единица из кода в СчX$_0$ при условии, что ТССХ находится в "0", то есть не окончена отработка строки. Подсчет необрабатываемых точек производится до нулевого состояния Сч X$_0$, о чем свидетельствует переключение триггера совпадения ТСХ$_0$ в "I". Вслед за этим восстанавливается код, характеризующий расстояние между местами обработки, передачей его из ΣX в СчX$_0$ и из ΣY в СчУ$_0$, и посылается сигнал на включение БФВП. Вслед за отработкой строки /ТССХ в "I"/ луч перемещается по оси У в результате добавления /или вычитания/ единиц в счетчик Сч У и одновременного вычитания единиц из кода в Сч У$_0$. По достижению заданного расстояния между строчками /ТСУ$_0$ в "I"/ пучок устанавливается в начальную точку очередной строки, благодаря передаче кода из Сч НХ в Сч Х, с одновременным восстановлением кодов в Сч X$_0$ и Сч У$_0$. Управление снова передается блоку БФВП. В такой последовательности операции управления обработкой и вычисление координат чередуются до тех пор, пока код в СчX не станет равным коду в Сч КХ и код в Сч У - коду в регистре РКУ, после чего вырабатывается сигнал "конец кадра", служащий признаком для выдачи из памяти /перфоввода/ следующей команды.

Алгоритмы отработки различных кадров даны на рис. 5.

Блок формирования временных параметров регулирует длительность импульсов и пауз между ними при любом их соотношении в диапазоне от 2 мксек до I0,2 сек. Число импульсов облучения каждой точки материала может задаваться в пределах от I до 2047. Предусмотрен также непрерывный режим обработки.

Для преобразования кодов в отклоняющий ток используется принцип суммирования на нагрузке взвешенных по двоичному за-

кону токов, формируемых стабилизаторами $Ст.1 \div Ст.10$ /рис. 6/ в соответствии с равенством

$$J_н = a_0 2^0 J_0 + a_1 2^1 J_0 + \cdots + a_n 2^n J_0 = J_0 \sum_{m=0}^{n} a_m 2^m$$

где a_m - 0 и 1, в зависимости от состояния m -го разряда счетчика отклонения.

С целью поддержания постоянного размера области обработки при изменении ускоряющего напряжения выходной ток преобразователя корректируется изменением опорного напряжения стабилизаторов. Реверсирование отклоняющих катушек осуществляется с помощью магнитоуправляемых лепестковых реле, имеющих высокое быстродействие и длительный срок службы.

Программное регулирование величины тока луча /сто возможных уровней, задается в процентах от номинального значения /и ускоряющего напряжения /16 ступеней, задается в киловольтах/ производится с помощью блока управления энергетическими параметрами /БУЭП/.

С основного или выносного пульта управления, являющегося частью блока преобразования и распределения информации /БПРИ/, программа вводится десятичными цифрами.

Отладка программ и контроль за правильностью их ввода облегчены благодаря наличию блока визуального контроля /БВК/, построенного на трубке с темновой записью и электролюминесцентной системе индикации, словами и десятичными цифрами отражающей практически все важные моменты управления технологическими процессами. В качестве иллюстрации возможности контроля отработки программы на рис. 7 показан скиатрон с выполненным на нем "Киевом-67" изображением".

Опыт эксплуатации показал, что описанная система управления способна обеспечить воспроизводимость электрических характеристик компонентов элионных интегральных схем, близкую к 100%.

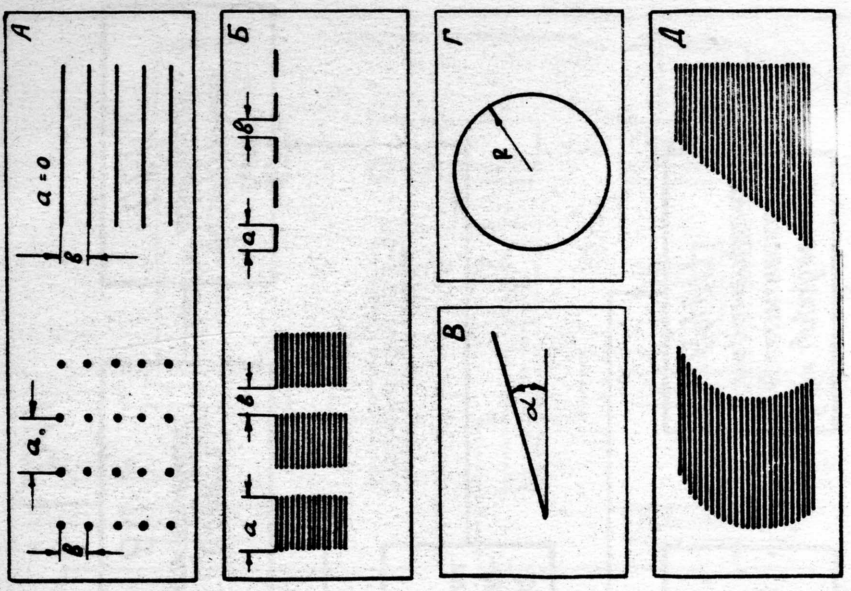


Рис. 1

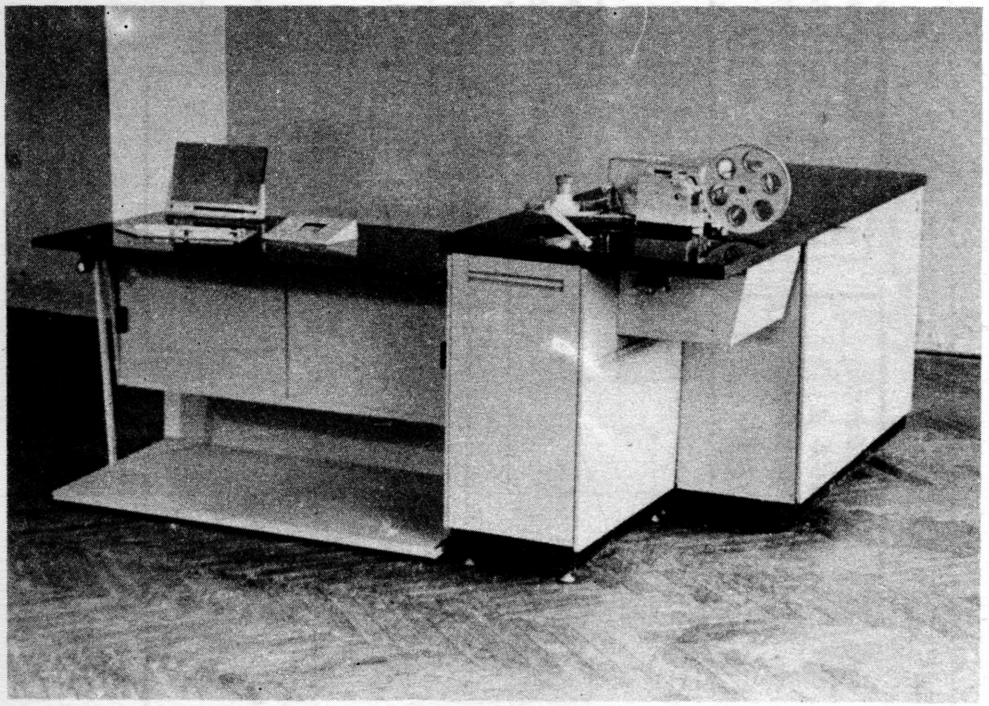


Рис. 2

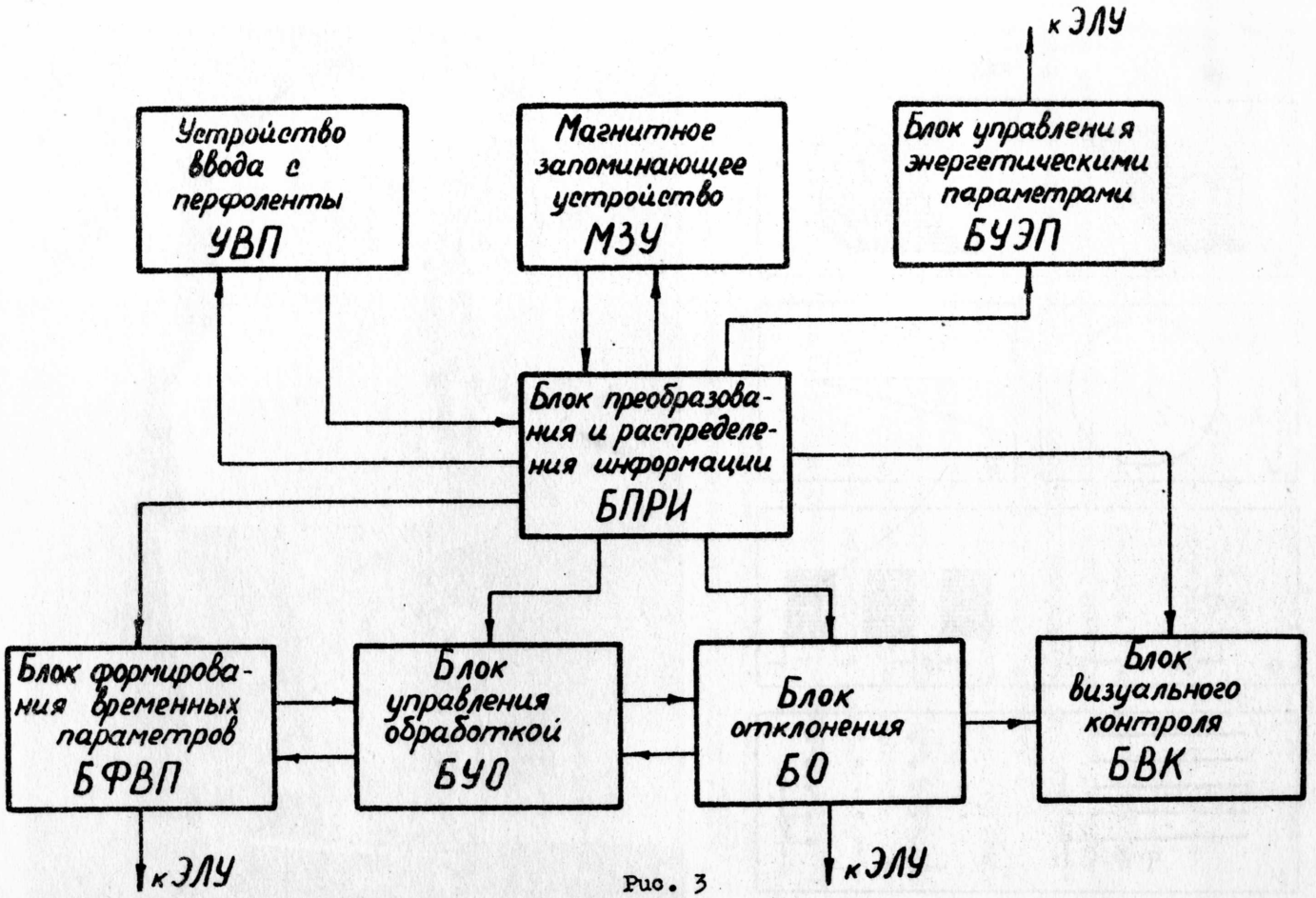

к ЭЛУ
Устройство ввода с перфоленты
УВП
Магнитное запоминающее устройство
МЗУ
Блок управления энергетическими параметрами
БУЭП
Блок преобразования и распределения информации
БПРИ
Блок формирования временных параметров
БФВП
Блок управления обработкой
БУО
Блок отклонения
БО
Блок визуального контроля
БВК
к ЭЛУ
к ЭЛУ

Рис. 3

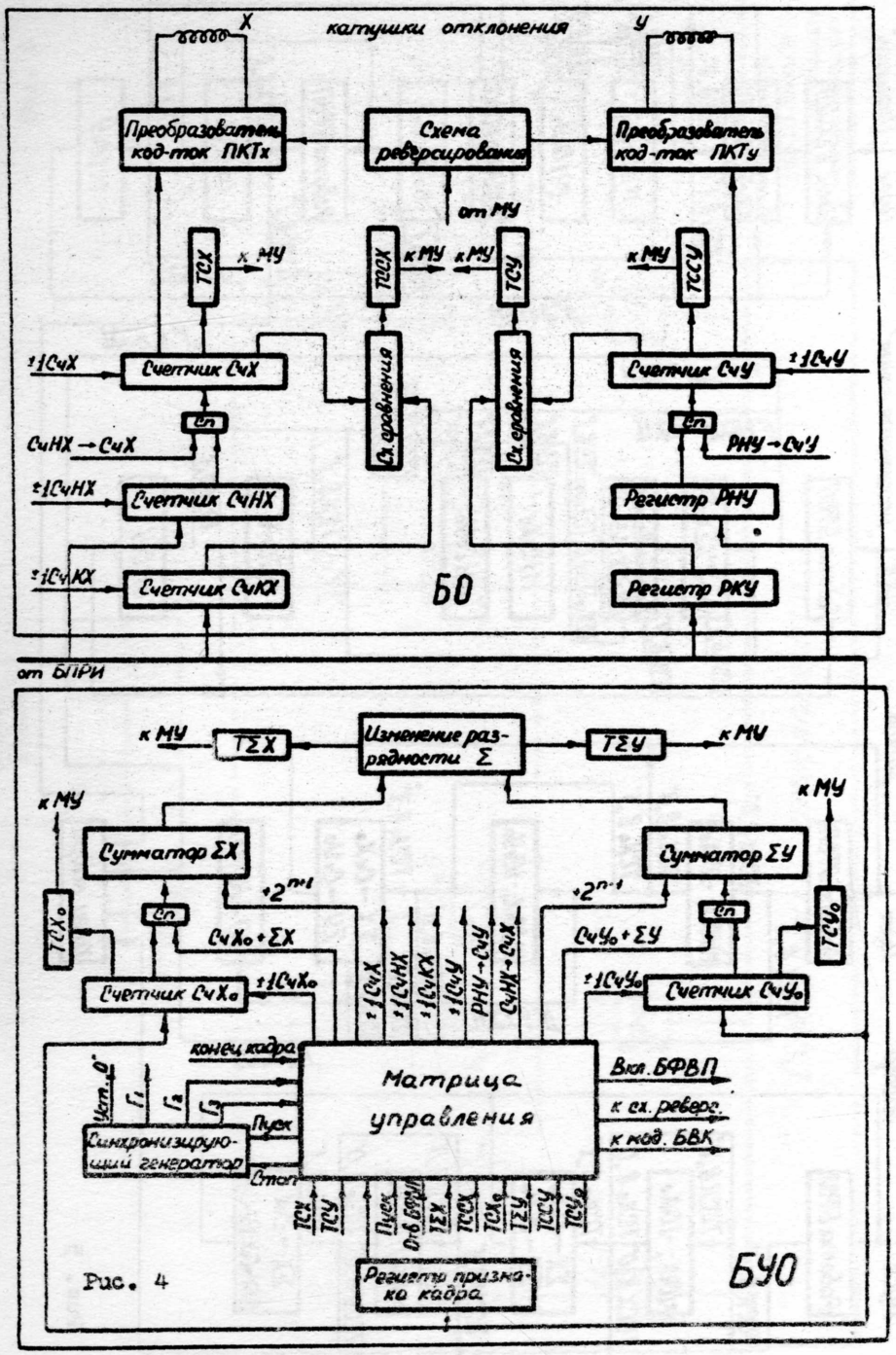

X
катушки отклонения
Y
Преобразователь код-ток ПКТх
Схема реверсирования
Преобразователь код-ток ПКТу
от МУ
ТСХ
к МУ
ТЗСХ
к МУ
к МУ
ТСУ
к МУ
ТЗСУ
±1СчХ
Счетчик СчХ
Сх. сравнения
Сх. сравнения
Счетчик СчУ
±1СчУ
Сп
СчНХ → СчХ
РНУ → СчУ
±1СчНХ
Счетчик СчНХ
Регистр РНУ
±1СчКХ
Счетчик СчКХ
БО
Регистр РКУ
от БПРИ
к МУ
ТΣХ
Изменение разрядности Σ
ТΣУ
к МУ
к МУ
к МУ
Сумматор ΣХ
Сумматор ΣУ
+2^m-1
+2^n-1
Сп
Сп
ТСХо
ТСУо
СчХо + ΣХ
СчУо + ΣУ
Счетчик СчХо
±1СчХо
±1СчУо
Счетчик СчУо
±1СчХ
±1СчНХ
±1СчКХ
±1СчУ
РНУ → СчУ
СчНХ → СчХ
конец кадра
Матрица управления
Вкл. БФВП
к сх. реверс.
к код. БВК
Уст. 0
Г1
Г2
Г3
Пуск
Синхронизирующий генератор
Стоп
ТСУ
ТСУ
Пуск об. БВУ
ТΣХ
ТЗСХ
ТСХо
ТΣУ
ТЗСУ
ТСУо
Регистр признака кадра
БУО

Рис. 4

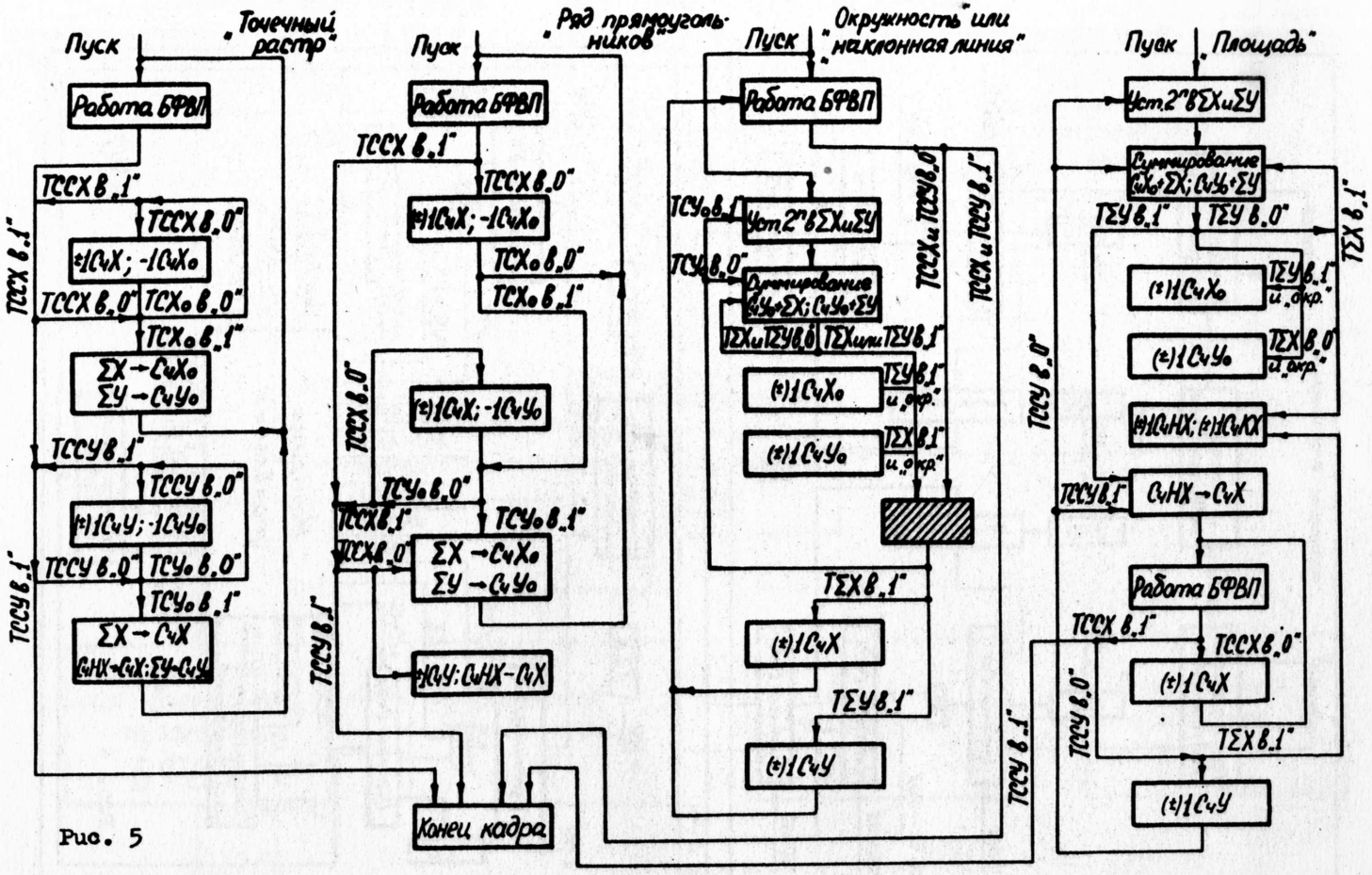
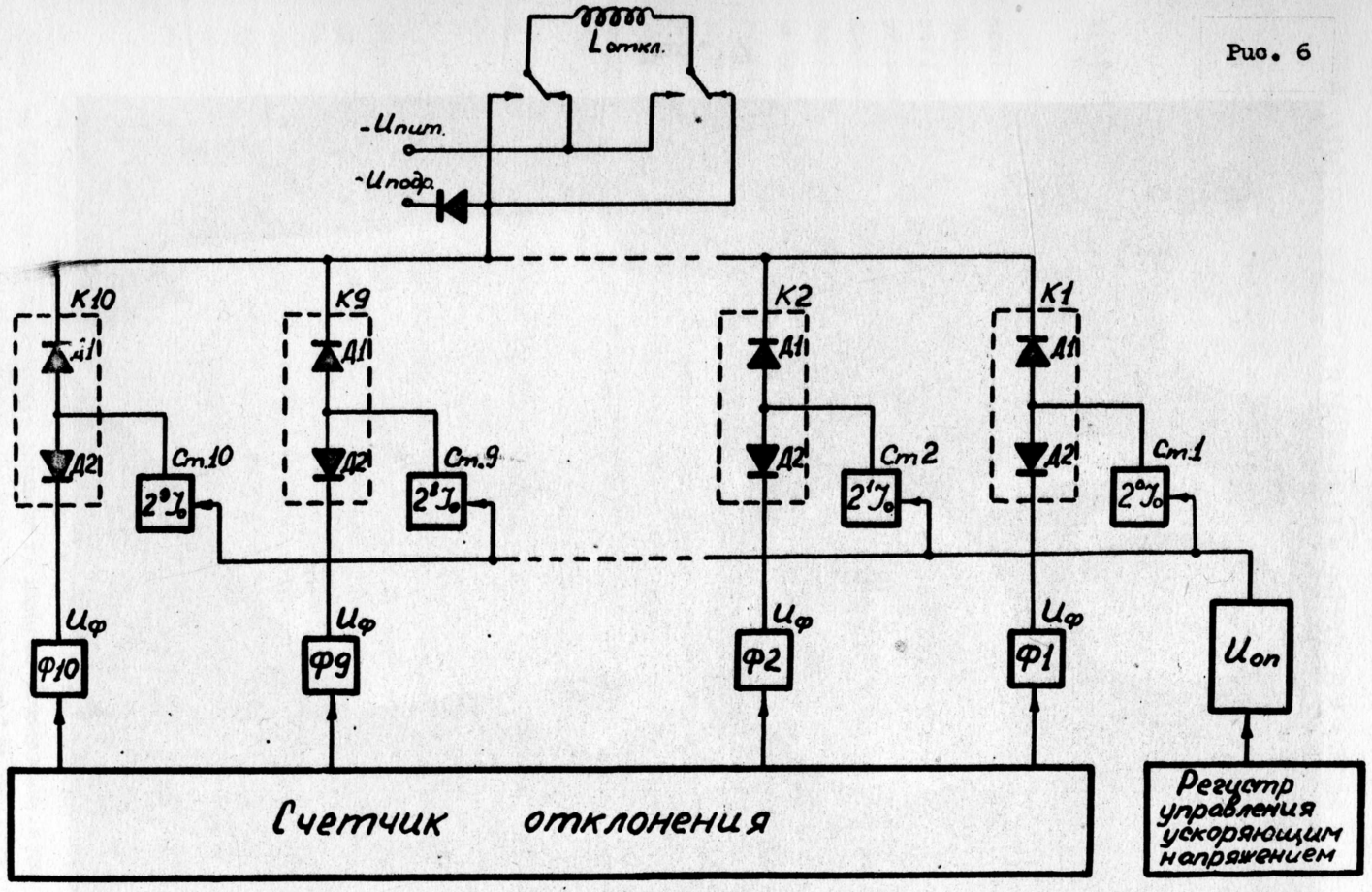

Пуск
"Точечный растр"
Работа БФВП
ТССХ в.1"
ТССХ в.0"
±1СчХ; -1СчХ0
ТССХ в.0"
ТСХ0 в.0"
ТСХ0 в.1"
ΣХ → СчХ0
ΣУ → СчУ0
ТССУ в.1"
ТССУ в.0"
±1СчУ; -1СчУ0
ТССУ в.0"
ТСУ0 в.0"
ТСУ0 в.1"
ΣХ → СчХ
СчНХ→СчХ; ΣУ→СчУ
ТССУ в.1"
Пуск
"Ряд прямоугольников"
Работа БФВП
ТССХ в.1"
ТССХ в.0"
±1СчХ; -1СчХ0
ТСХ0 в.0"
ТСХ0 в.1"
ТССХ в.0"
±1СчХ; -1СчУ0
ТССУ в.0"
ТСУ0 в.0"
ТССХ в.1"
ТССХ в.0"
ТСУ0 в.1"
ΣХ → СчХ0
ΣУ → СчУ0
ТССУ в.1"
±1СчУ; СчНХ→СчХ
Пуск
"Окружность" или "наклонная линия"
Работа БФВП
ТСУ0 в.1"
Уст.2^nвΣХиΣУ
ТСУ0 в.0"
Суммирование СчХ0+ΣХ; СчУ0+ΣУ
ТССХи ТССУ в.0"
ТССХ и ТССУ в.1"
ТΣХиТΣУ в.0"
ТΣХ или ТΣУ в.1"
(±)1СчХ0
ТΣУ в.1" и "окр."
(±)1СчУ0
ТΣХ в.1" и "окр."
ТΣХ в.1"
(±)1СчХ
ТΣУ в.1"
(±)1СчУ
Пуск
"Площадь"
Уст.2^nвΣХиΣУ
Суммирование СчХ0+ΣХ; СчУ0+ΣУ
ТΣУ в.1"
ТΣУ в.0"
ТΣХ в.1"
(±)1СчХ0
ТΣУ в.1" и "окр."
(±)1СчУ0
ТΣХ в.0" и "окр."
±1СчНХ; (±)1СчКХ
ТССУ в.0"
ТССУ в.1"
СчНХ→СчХ
Работа БФВП
ТССХ в.1"
ТССХ в.0"
(±)1СчХ
ТΣХ в.1"
ТССУ в.0"
(±)1СчУ
Конец кадра
ТССУ в.1"

Рис. 5

Рис. 6



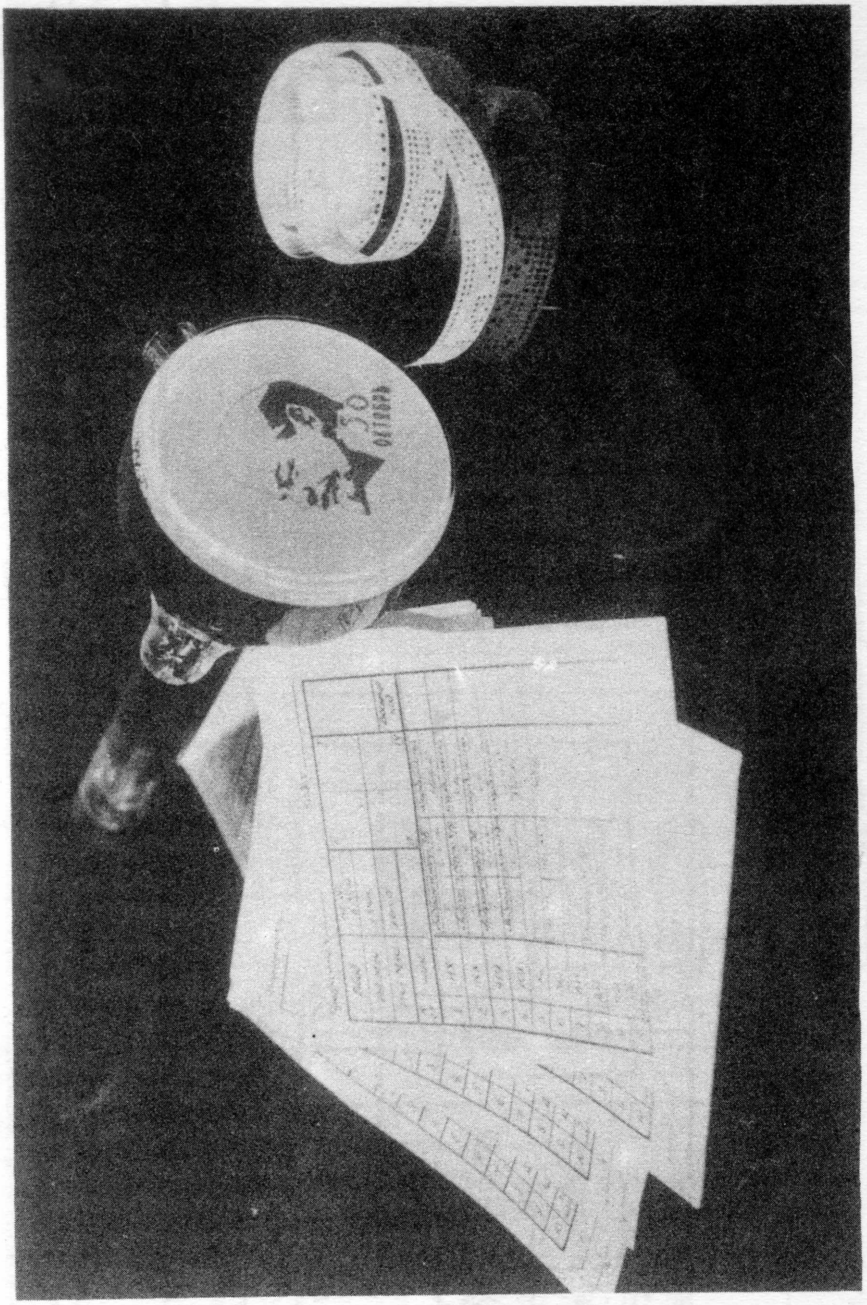


Рис. 7

A MICRO PATTERN POSITIONING SYSTEM

Y. Oshima and B. S. Chang
Institute of Industrial Science, Univ. of Tokyo
Tokyo, Japan

1. INTRODUCTION

The micro pattern positioning system is necessary for automatic assembling of transistors and integrating circuits[1]. This newly developed micro pattern positioning system aims to automatize the wire bonding process for transistor pellets, in which positioning is done manually at the present. Positioning of pellets in the transistor assembling process is done in this country by young woman workers by means of stereoscopic microscopes and motion reducing mechanisms. The wire bonding process is becoming the neck of transistor mass production because of shortage of young woman workers. The problem of automatization of pellet positioning is the optical detection system which replaces the human eye. In this automatic positioning system the special photoelectric microscope with L-type slits is used. This paper describes the principle and construction of the micro pattern positioning system as well as the experimental results of the prototype system.

2. OPTICAL PATTERN DETECTION SYSTEM

2.1 Micro pattern

If the planar or mesa type transistor pellets are illuminated from the direction with the angle of 45° to the vertical, the optical patterns with good contrast as shown in Fig 1 are obtained. The polished specular surface of the semiconductor material such as Si appears black, while the surface of Al deposited electrodes appears brighter than the background because the electrode surface reflects light diffusely and some of the light is reflected in the vertical direction. Such optical patterns are the object of position detection.

2.2 Principle of optical pattern detection system

The system uses the photoelectric microscope[2] as the detector. As shown in Fig 2, the optical pattern obtained by inclined illumination is enlarged by the optical system and its real images are formed at two slits

after separation of the optical beam into two directions by means of the half-silvered mirror. Two photomultipliers receive the light through the slits. Since two slits are shifted relatively with respect to the optical axis, the outputs from the photomultipliers are shifted to each other. These output signals are introduced to the differential amplifier which amplifies the difference. The output of the differential amplifier is applied to the pulse shaper and then differentiated. Thus the recognition pulse signal is generated at the precise position of the optical pattern.
Fig 3 shows the wave forms of each part. In this photoelectric microscope the slit width and relative shift of two slits are important factors. Let us consider the case of rectangular optical pattern for simplicity. Figs 4 and 5 show the schematic output wave forms in cases of Wp = Ws and Wp $>$ Ws, where Wp is the pattern width and Ws is the slit width. In case of Wp = Ws the allowable range of the relative shift of slits is 0 ~ 2 Wp and the output wave form is triangular. In case of Wp $>$ Ws the allowable range is 0 ~ (Wp + Ws) and the output wave form is trapezoidal. The former is the special case of the latter. Three values of the relative shift of slits $(Wp - \frac{Ws}{2})$, Wp and $(Wp + \frac{Ws}{2})$ are chosen and the experiments determine which is the best as described below.

In order to determine the position in X and Y directions, slits are arranged as shown in Fig 6. S_{x1} and S_{x2} are the slits for X direction and S_{y1} and S_{y2} are the slits for Y direction. We call this arrangement as L-type slit. When the optical pattern is fed in the Y direction, the outputs of two photomultipliers P_1 and P_2 as shown by y_1 and y_2 in the figure are obtained and the recognition pulse p_y is generated. The optical pattern is still fed by the predetermined distance from the point of generation of p_y and is stopped. Then the optical pattern is fed in the X direction. When it is brought to the point corresponding to the slits S_{x1} and S_{x2}, the outputs of photomultipliers P_1 and P_2 as shown by x_1 and x_2 in the figure are obtained and the recognition pulse p_x is generated. The otpical pattern is stopped at the instant of generation of p_x. Thus the optical pattern is positioned at the predetermined point.

Fig 7 shows the optical system. In order to enlarge the pattern the 5X microscope objective is used. The distance between the objective lens and the object is chosen as 50mm so as to enable the bonding work. The optical path is divided into two by half-silvered mirror m_1. The real images of the pattern are formed at the slits S_1 and S_2. The photomultipliers P_1 and P_2 collect the light through S_1 and S_2 respectively and convert the

optical signal to the electrical signal. The half-silvered mirrors m_2 and m_3 reflect a portion of light to the eyepieces. The eyepiece is a 10X lens. The operator can see the pattern through the eyepiece with magnification of 50. The range of adjustable slit width is 0~2 mm and the range of adjustable position of each slit is ±2 mm. Both adjustments are done by means of respective micrometer screw. Fig.8 shows the outside view of the above-mentioned photoelectric microscope together with the driving mechanism to be described in the following chapter.

3. DRIVING MECHANISM

The transistor pellet to be assembled is placed on the small table which is driven in both X and Y directions by means of micrometer screws. The driving mechanism of Y direction consists of the stepping motor for fine positioning, induction motor for quick feed, magnetic clutch with brake for switching of quick feed and stepping motor feed, gear train and micrometer screw. For X direction, the induction motor and magnetic clutch are not necessary. The quick feed by induction motor is 5 mm/sec. Since the gear ratio and screw feed for stepping motor drive is so chosen that one step rotation of stepping motor results 2 μ linear motion and the maximum pulse frequency for the used stepping motor to follow up is 200 pulses per sec., the maximum stepping motor feed is 400 μ/sec. The stroke of table feed is 20 cm in both X and Y directions.

4. CONTROL CIRCUITS

The control circuits consist of the control counters, gate inverters, flip-flops, monostable multivibrators, multivibrator oscillator and various driving circuits. Fig 9 shows the block diagram of control circuits. Fig 10 shows the timing chart of the control circuits. The function of the circuits is as follows:

At first setting RSA "1" and pushing the push button B, the output of the monostable multivibrator MMA resets all the packages. Then setting RSA "2" and pushing the push button B, the output of MMA sets the flip-flop FFC. Setting of FFC through the clutch driving circuit CD activates the magnetic clutch and the quick feed in Y direction starts. When the table is fed in Y direction and the optical mark on the table reaches just under the optical detection system, the output pulse of the detection system is generated and is introduced to the input terminal of the control circuits.

When the control counter A consisting of flip-flops FFA and FFB counts "1", the state of the gate inverter GIA changes from "1" to "0", resetting FFC and setting FFD. Thus the magnetic clutch is deenergized and the induction motor stops, while the stepping motor starts. When the table is driven by the stepping motor and the optical pattern reaches the point corresponding to S_y slits of the optical detection system, the detection system generates the recognition pulse which is introduced to the control circuits. When the control counter A counts "2", the state of the gate inverter GIB changes from "1" to "0", letting the control counter B start counting. The stepping motor feed still continues. The control counter B counts the driving pulses of stepping motor generated by the multivibrator oscillator MO. When the control counter B counts the number preset by the dials, the output pulse is generated, resetting FFD and setting FFE. Resetting of FFD stops the Y-direction stepping motor feed.

Setting of FFE starts the stepping motor feed in X-direction. When the optical pattern reaches the point corresponding to S_x slits of the optical detection system, the detection system generates the recognition pulse. When the control counter A counts "3", the state of the gate inverter GIC changes from "1" to "0", resetting FFE and setting FFF. Resetting of FFE stops the X-direction stepping motor feed. Thus the positioning action is completed. Setting of FFF generates the command to start the bonding operation. The change of state of GIC through the monostable multivibrator MMC resets all the packages. When the bonding operation is completed, the signal is generated and the above-mentioned actions are repeated. Each action is indicated by the lamps LA~LD. Fig 11 shows the flow chart of sequential action. Fig 12 shows the block diagram of the overall system.

5. EXPERIMENTS

Various experiments have been done concerning the prototype system. The experimental items are the effects of pellet position, slit width and relative shift of slits, pellet type, illumination and rotational displacement of pellet. The experimental results of each item are to be described below. Before positioning experiments, the characteristics of driving mechanism were tested. In this test the small table is fed by pulse number preset on the control counter B by means of the stepping motor. The actual displacement of the table is measured by the micrometer microscope equipped in the eyepiece of photoelectric microscope. The measured displace-

ment and calculated displacement based on pulse number are compared. The measured data are the average of 5 measurements. Three kinds of feed rate 100 μ/sec, 200 μ/sec and 400 μ/sec are taken. Fig 13 shows the test results. Fig 13 (a) shows that the error increases with number of steps owing to the accumulated pitch error of micrometer screw. This error is not important for positioning accuracy. Fig 13 (b) shows the maximum deviation among 5 measured data at each step. This deviation is within $\pm 3\mu$ in the range up to 600 steps. This error affects the positioning accuracy but doesn't depend upon feed rate.

5.1 Effects of pellet position

The pellet position means the position of optical pattern in X-direction in case of Y-direction feed and the position in Y-direction in case of X-direction feed. The pellet position is measured from the center of optical field. The effects of pellet position are investigated every 100μ displacement up to 1 mm. Figs 14 (1) ~ (3) show the experimental results. In these figures (1), (2) and (3) are the cases of pellet type 2 SC 455, 2 SC 454 and 2 SC 281 as shown in Fig 1 respectively. The experimental conditions are:

Supply voltage for illuminating lamp : 7 V
(SG-2, 8 V, 60 W)
Direction of illumination : opposite to feed direction
Feed rate : 400 μ/sec

The results show that the positioning error increases with pellet position especially in case of slit width = 300μ expressed in terms of dimensions of pattern (actual width is 5 times as much as this value because of magnification of optical system). The main reason is assumed that the intensity of reflected light from the pellet varies in accordance with its position and the gain characteristics for amplifying two photomultiplier outputs are unbalanced. If two shifted output wave forms of photomultipliers vary in the same way according to the pellet position as shown in Fig 15(a), no error occurs. But if two wave forms vary in different way as shown in Fig 15 (b), the positioning error occurs. Fig 16 shows the output wave forms of photomultipliers. As understood from Fig 14, the effects of pellet position is very small if the slit width is small.

5.2 Effects of slit width and relative shift of slits

Taking the slit widths at intervals of 50μ from 50μ to 300μ, the positioning accuracy is measured. The three kinds of relative shift of slits as

described in Sec. 2.2 are chosen. The other conditions are the same with those described in Sec. 5.1. Figs 17(1)~(3) show the experimental results. In these figures, (1), (2) and (3) are the cases of pellet type 2SC 455, 2SC 454 and 2SC 281. The results show that the error decreases with decrease of slit width except the case of 2SC 281 pellet as shown in Fig 17(3). In case of 2SC 281 pellet which has level change on the surface, the brilliant small portion appears near the electrode pattern as shown in Fig 1(c). This is due to level change. In this case the large error occurs when the slit width is very small as shown in Fig 17(3). Considering such a case and other noise effects the optimum slit width is taken as 60 ± 5% of the electrode pattern length in feed direction. If there is no noise effect, the smaller slit width results the smaller error range as shown in Fig 18 because the output wave forms of photomultipliers become narrow.

As recognized from Fig 17, in almost every case the best results are obtained when the relative shift is chosen as $(Wp + \frac{Ws}{2})$. The larger relative shift of slits results the smaller error range as shown in Fig 19. Fig 20 shows the output wave forms of photomultipliers in cases of various slit widths and two pellet positions. Fig 21 shows the case of each relative shift of slits.

5.3 Effects of pellet type

Rearranging the experimental results as shown in Figs 17(1)~(3) with respect to pellet type, the results as shown in Fig 22 are obtained. As understood from the figures, the error is of the same order for each pattern with almost same size when the same slit width is used. The effects of pellet type are conspicuous when the small slit width and small relative shift of slits are used or when there exists the rotational displacement of pellet as described below.

5.4 Effects of illumination

Choosing the supply voltage for illuminating lamp as 6, 7 and 8 V, experiments are done. The other conditions are:

Slit width : 200 μ
Relative shift of slits : $Wp + \frac{1}{2} Ws$
Feed rate : 400 μ/sec
Direction of illumination : 180° (opposite) and 90° to feed direction.

Fig 23 shows the experimental results. The results show that the effects of intensity of illumination are very small. Next the effects of direction of

illumination are investigated. Two kinds of direction 90° and 180° with respect to the direction of feed are taken. The results are indicated for each pellet type as shown in Fig 24. Rearranging the results we obtain Fig 25. As understood from the figure, in cases of pellet types 2SC 454 and 2SC 455 the illumination from 90° direction in Y-direction feed and that from 180° direction in X-direction feed are preferable but in case of pellet type 2SC 281 the inverse condition is preferable.

5.5 Effects of rotational displacement of pellet

Experiments are done at rotational positions of pellet -20°, -10°, 0°, +10° and +20° and at pellet position of 200μ distance from the center, choosing the slit width as 100, 150, 200, 250 and 300 μ. Fig 26 shows the experimental results. The effects of rotational displacement of pellet are conspicuous. In case of pellet type 2SC 455 the larger slit width gives the smaller positioning error for small rotational displacement of pellet. The pellet type strong for rotational displacement is in order of 2SC 281, 2SC 454 and 2SC 455.

6. POSITIONING FUNCTION

Positioning function is estimated by positioning accuracy, positioning time and range of allowable pellet position both rotational and linear. Some interconnection exists among them. For example, if we specify the allowable pellet position in narrow range, good positioning accuracy is obtained. On the contrary, if we extend the range of pellet position, bad positioning accuracy or in the worst case failure of positioning results.

6.1 Positioning accuracy

Positioning accuracy depends upon the initial condition of pellet, that is, linear or rotational pellet position. The positioning accuracy of the prototype system is tabulated in Table 1. This shows that the accuracy ±5μ is obtained if the pellet position is within 200μ and the rotational displacement is less than ±10°. Fig 27 shows the error distribution range on the enlarged pattern on the following conditions:

Pellet position in X direction (Y-direction feed)	:	1,000 μ
Pellet position in Y-direction (X-direction feed)	:	200 μ (in actual positioning action this is negligibly small)
Rotational displacement of pellet	:	± 20°

This figure clarifies the positioning situation.

6.2 Positioning time

Positioning time depends upon the initial condition of pellet and feed rate. Designating symbols of each dimension as shown in Fig 28, we obtain the positioning time as

$$T = A_y/V_f + (a_y + b_y + c_y + d_y + a_x + \frac{1}{2}b_x + \frac{1}{2}c_x)/Vs$$

where V_f is quick feed rate and Vs is stepping motor feed rate.
Taking the pellet type 2SC 281 as an example,

$b_y = 640\,\mu$, $C_y = 340\,\mu$

$b_x = 580\,\mu$, $C_x = 280\,\mu$

are given, where b_x and b_y are determined by positioning conditions.
Assuming $A_y = 5$mm, $a_y = 20\mu$, $d_y = 20\mu$, $a_x = 500\,\mu$,

$V_f = 5$ mm/sec, $Vs = 400\mu$/sec

we obtain $T = 5.87$ sec

Table 2 shows the calculated positioning time for each type of pellet. In total average the calculated positioning time is 5.13 sec in case of initial position = 5mm. This time is not satisfactory. The problem will be solved by using a more high speed stepping motor.

6.3 Range of allowable pellet position

The range of allowable pellet position is determined by the specified positioning accuracy. Table 3 shows such a range for each type of pellet. This range specifies the accuracy of the previous processes which the pellets undergo before they are introduced to the positiotioning system.

7. ANALYSIS OF POSITIONING ERROR

The error sources and order of each error are investigated. In this positioning system positioning is done by stepping motor feed, that is, stepwise feed (1 step = 2μ). The recognition pulse generated anywhere in a step motion stops the pellet at the end of this step motion. The maximum error from such a source which is designated ε_{m1} is equal to 1/2 step. That is $\varepsilon_{m1} = \pm 1\mu$. Then the overshoot of stepping motor due to inertia is considered. This error designated ε_{m2} is equal to 1 step and always positive. Therefore $\varepsilon_{m2} = 2\mu$. Next the mechanical errors such as backlash and pitch error of micrometer screw cause the positioning error ε_{m3}. As described in Chap 5, $\varepsilon_{m3} = \pm 3\mu$. Thus the positioning error due to driving mechanism is ($\varepsilon_{m1} + \varepsilon_{m2} + \varepsilon_{m3}$).

In the second place the error caused from the detection system is investigated. This error consists of the error due to differential amplifier itself and the error due to the unbalance of output wave forms of photomultipliers generated by pellet pattern and also the difference of input from the back ground of optical pattern. The error due to differential amplifier itself, ε_{d1}, is measured as $\pm 1\mu$. That is $\varepsilon_{d1} = \pm 1\mu$. The other error is illustrated in Fig 29. In the figure (a) is the normal case, (b) is the unbalance of output wave forms, and (c) is the difference of input from the back ground. The unbalance of output wave forms including the difference of back ground is caused from the pellet position, rotational displacement and other noises. The positioning error due to the unbalance of output wave forms is designated ε_{d2}.

The maximum positioning error measured is $\pm 15\mu$. That is $(\varepsilon_{m1}+\varepsilon_{m2}+\varepsilon_{m3}) + (\varepsilon_{d1}+\varepsilon_{d2}) = \pm 15\mu$

Therefore $\varepsilon_{d2} = \pm 12\mu$

Since the error due to the pellet position is within $\pm 2\mu$ if the suitable slit width is chosen, the most part of ε_{d2} is due to the rotational displacement of pellet.

8. Remarks and suggestion for application

In order to improve the positioning accuracy it is necessary to increase the accuracy of detection system. For this purpose the noise should be avoided. The rotational displacement of pellet is considered as a kind of noise to the detection system and is the most important one. In order to improve the accuracy it is necessary to limit the rotational displacement within narrow range or to design the pellet pattern strong for rotation. For the former purpose the previous processes should be strictly controlled.

Requirements for the pellet is:

1) Symmetrical pattern whose contour is similar to circular arc as much as possible is preferable.

2) The arrangement of pattern in the pellet as shown in Fig 30 is good for illumination. This arrangement is conventional.

3) The pellet whose electrode surface has uniform infinitesimal roughness and whose semi-conductor surface is mirror finished is preferable. The method of loading the transistor pellet to the positioning system is a problem. In loading the arrangement as shown in Fig 31 is suggested.

If such an arrangement is adopted, continuous positioning is possible and positioning time can be saved sufficiently.

9. CONCLUSION

The principle and construction of a newly developed micro pattern positioning system for automatic assembling of transistors as well as the experimental results concerning the characteristics of the prototype have been described. The experimental results show that the positioning with accuracy $\pm 5\mu$ is possible under a certain condition which is realizable in the previous processes. If the positioning time is improved by using high speed stepping motor, this sytem can be put to practical use.

Appreciation is expressed to Mr. T. Numakura and Mr. K. Isoda, Hitachi Central Lab., for their active co-operation in the development of this system.

REFERENCES

1) R. L. MOORE, High-speed servo positioner bonds mesa transistors, Electronics, Feb. 8, 1963, 58 - 61.

2) S. NISHIDA, Y. DOI, and K. TOGINO, Numerical control of jig borer 'Jidic', Proc. 1st Congr. I.F.A.C. Moscow, IV (1961) 42 - 47, Butterworths.

(a) Planar type
2SC 455
($320\mu \times 160\mu$)

(b) Planar type
2SC 454
($270\mu \times 190\mu$)

(c) Mesa type
2SC 281
($340\mu \times 280\mu$)

Fig 1 Optical patterns

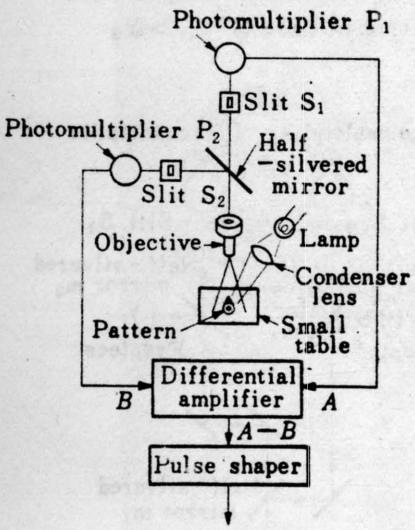


Fig 2 Principle of photoelectric microscope

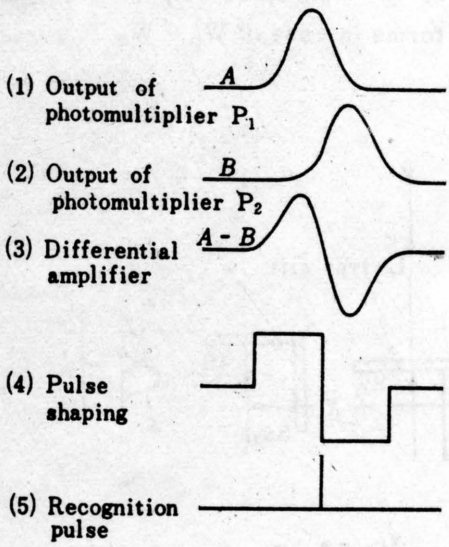


Fig 3 Wave forms of each part

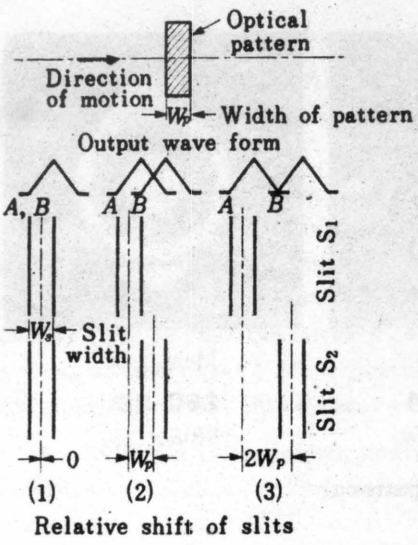


Fig 4 Schematic output wave forms in case of $W_p = W_s$

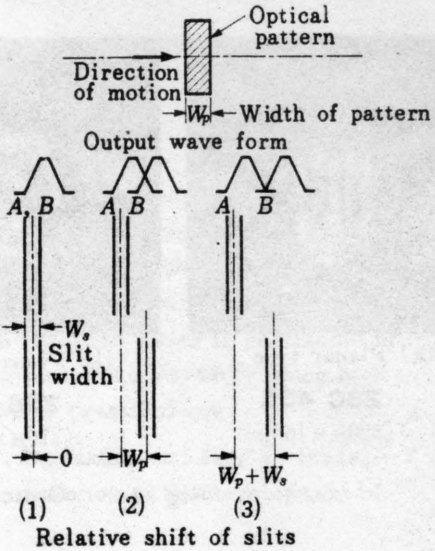


Fig 5 Schematic output wave forms in case of $W_p > W_s$

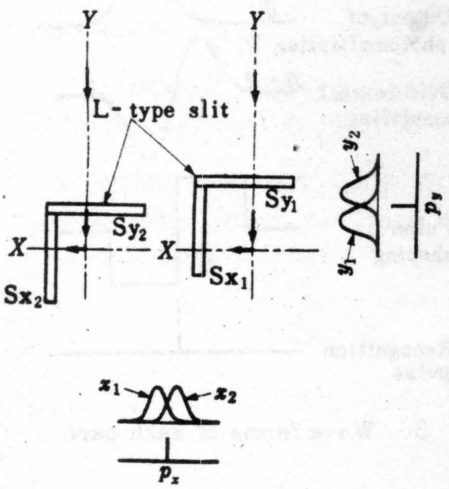


Fig 6 L-type slit

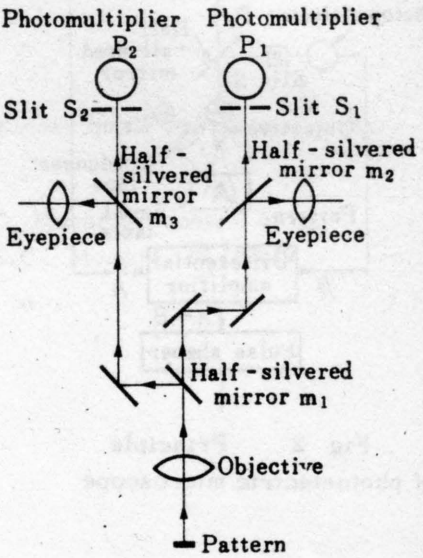


Fig 7 Optical system

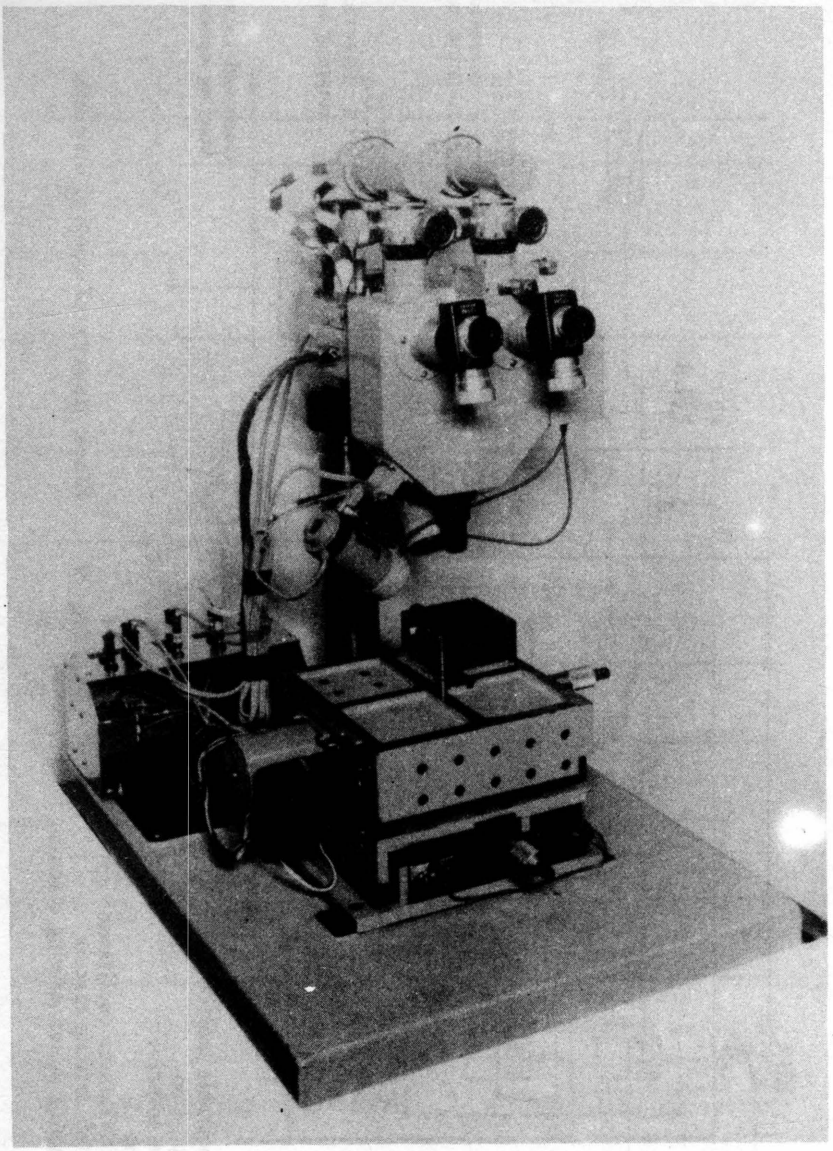
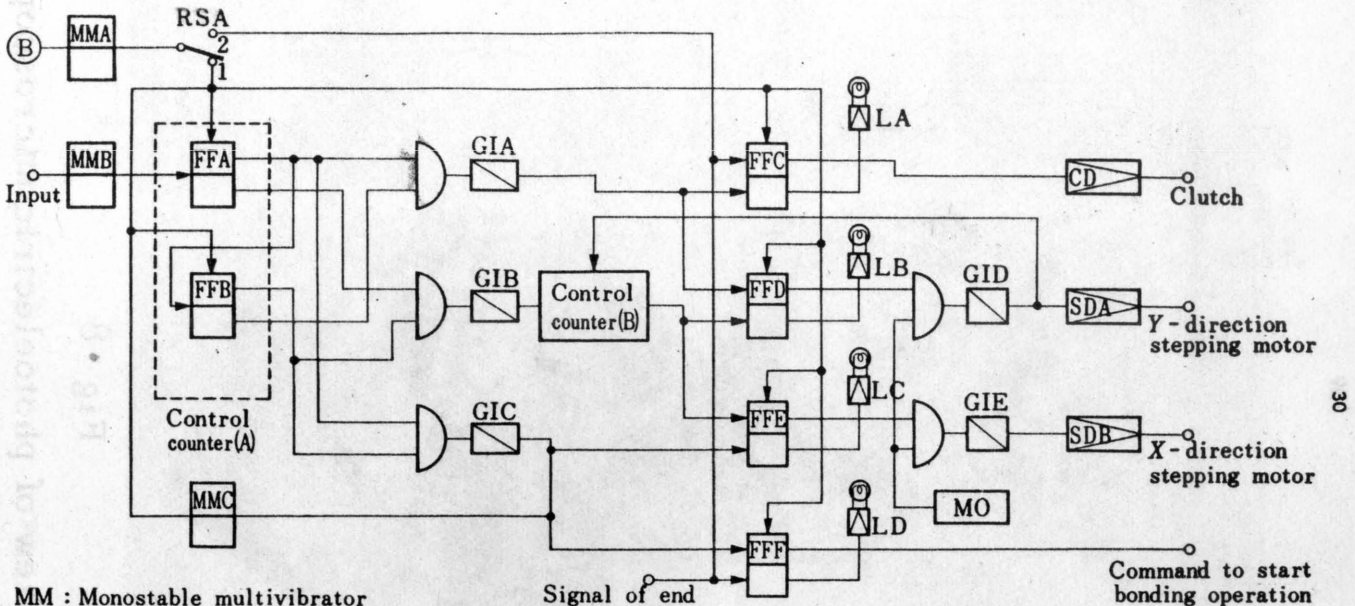


Fig 8

Outside view of photoelectric microscope



MM : Monostable multivibrator
FF : Flip-flop
GI : Gate inverter
L : Lamp
Mo : Multivibrator oscillator
CD : Clutch driving circuit
SD : Stepping motor driving circuit

Fig 9 Block diagram of control circuits

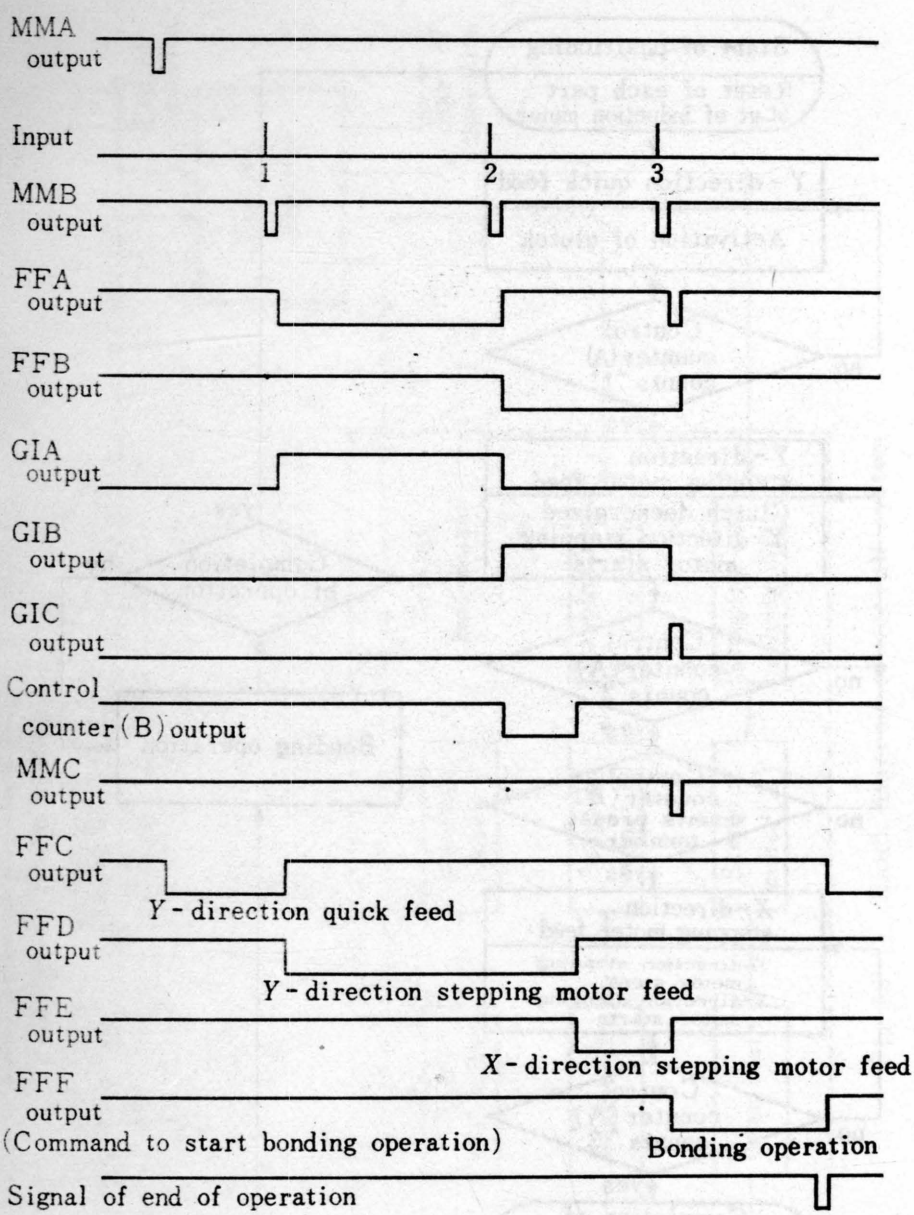


Fig 10 Timing chart of control circuits

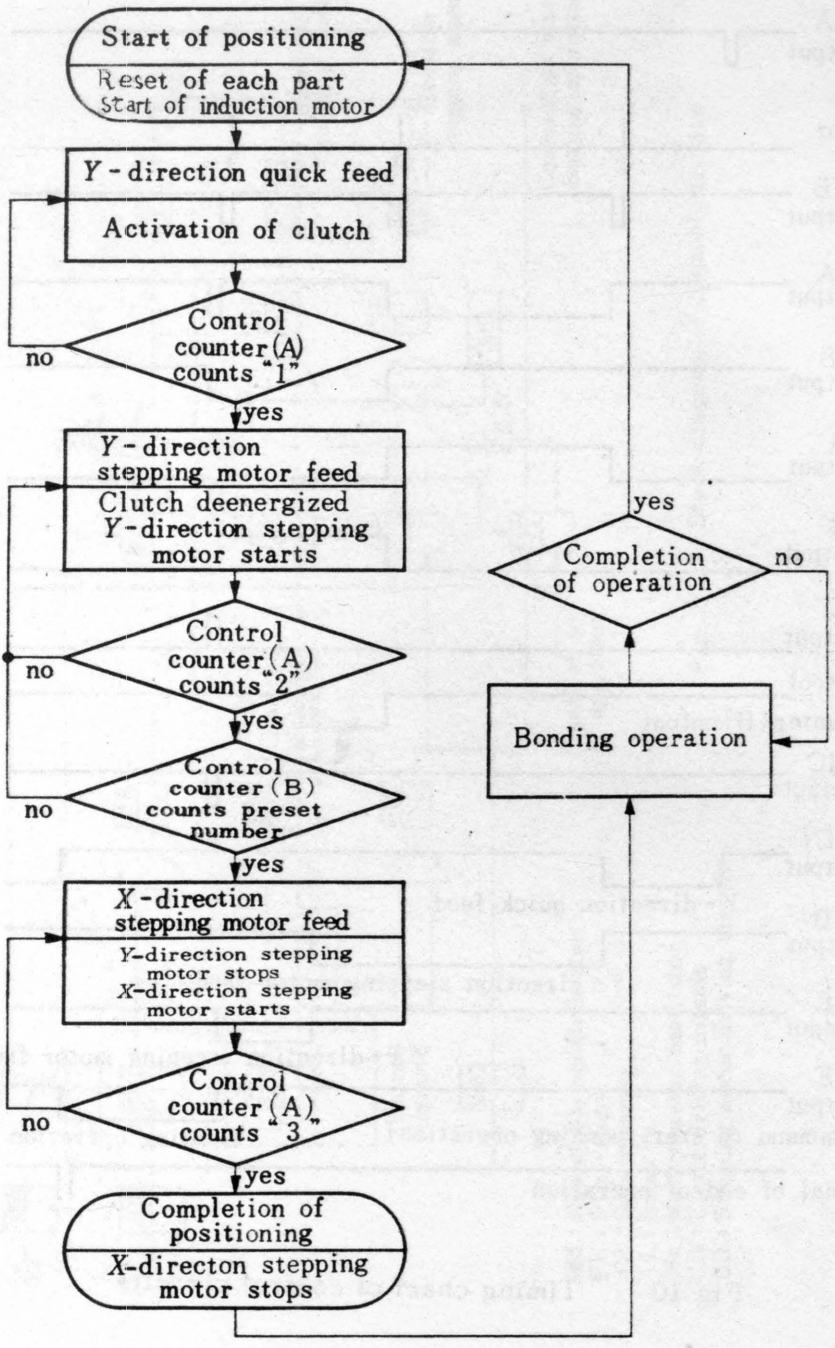


Fig 11 Flow chart of sequencial action

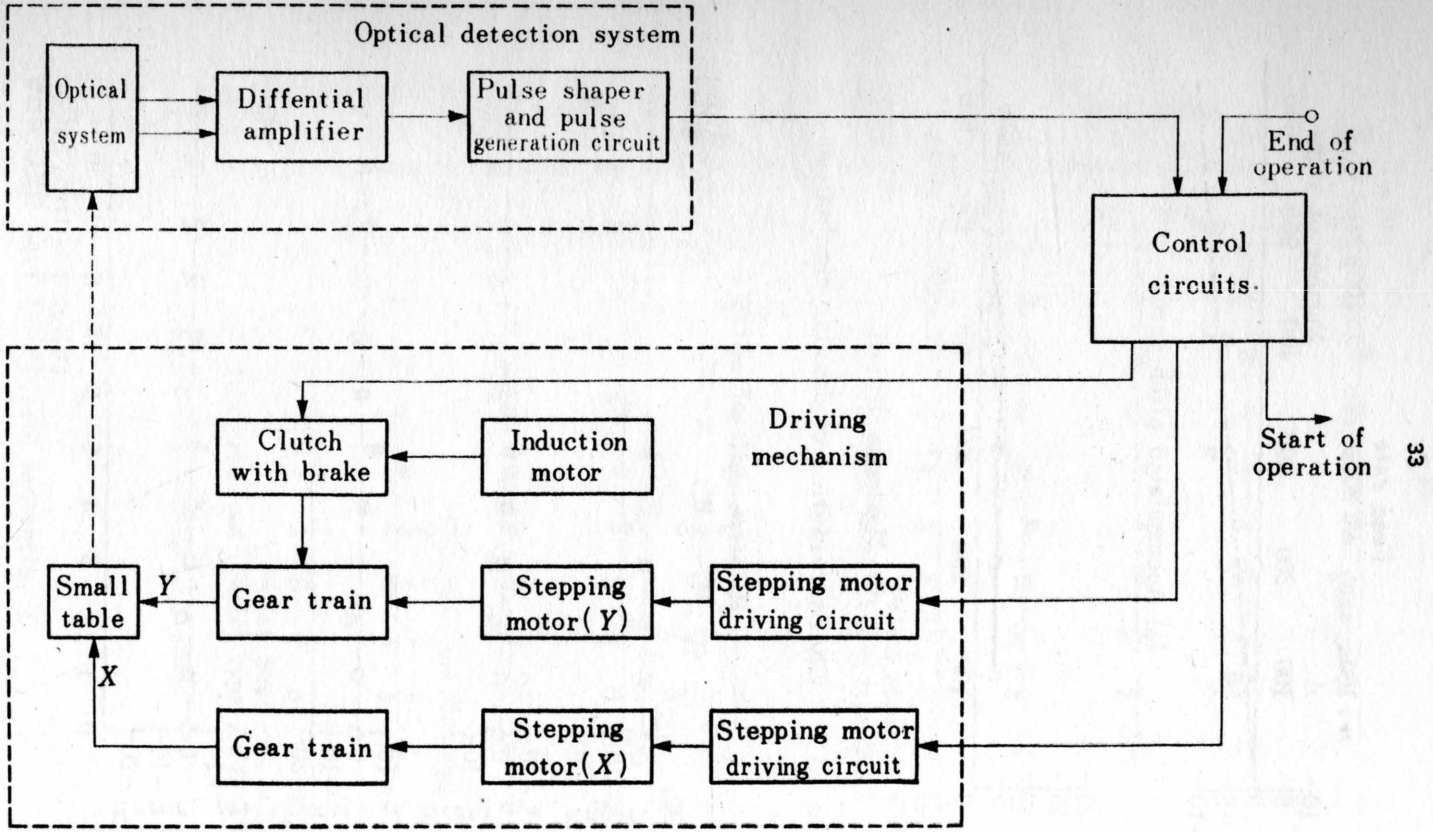
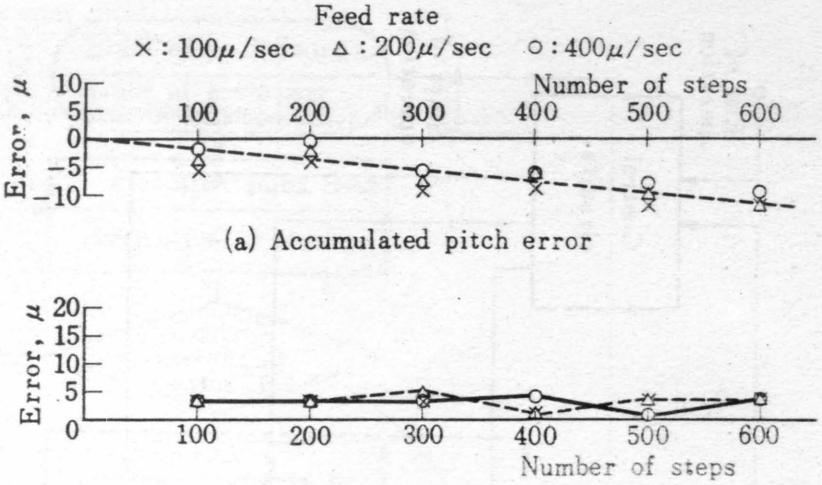


Fig 12 Block diagram of overall system



(a) Accumulated pitch error

(b) Maximum deviation

Fig 13 Characteristics of driving mechanism

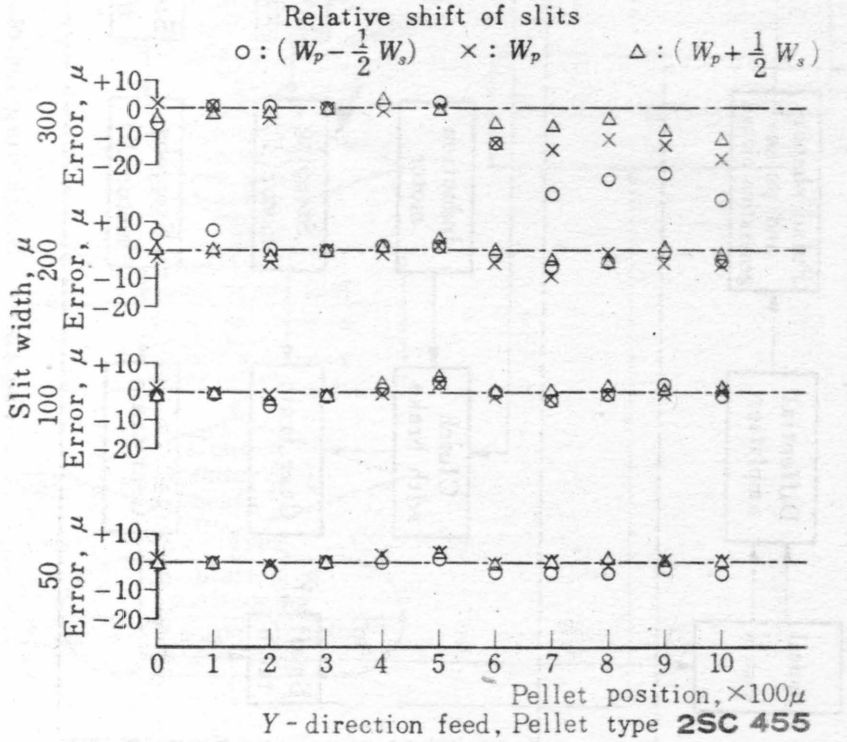
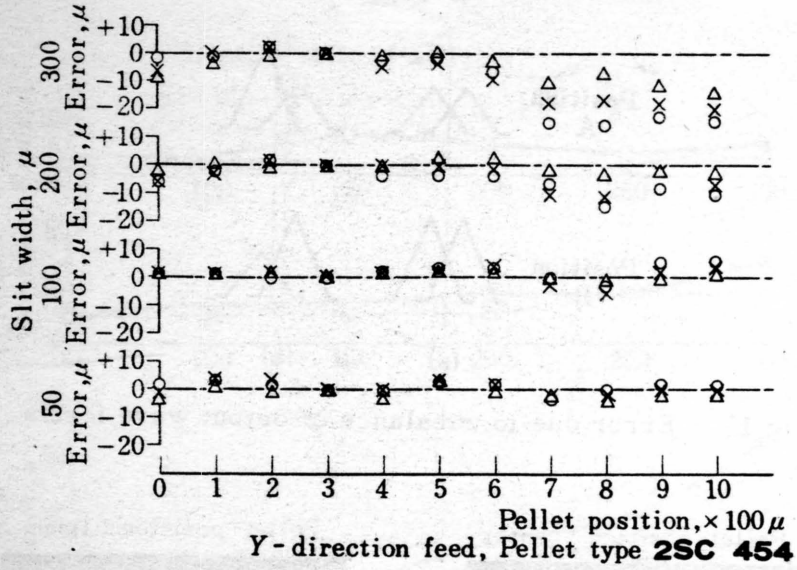


Fig 14(1) Effects of pellet position



(2)

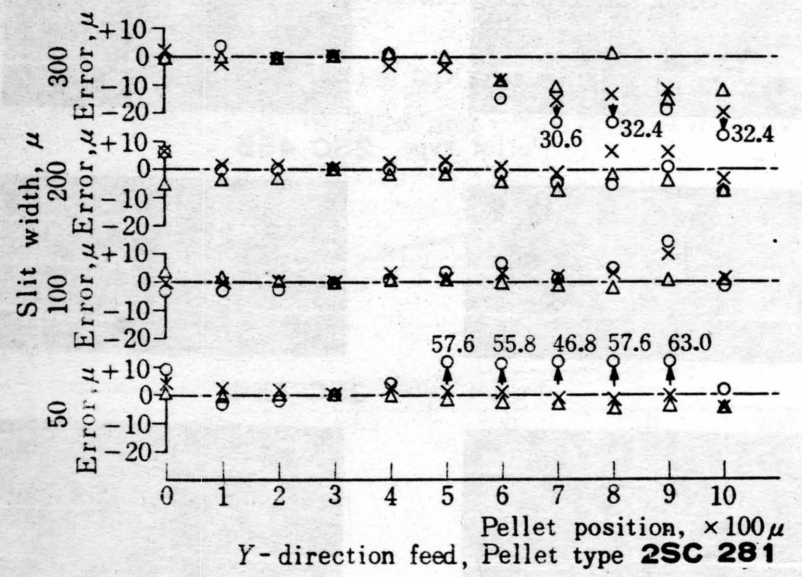


Fig 14(3) Effects of pellet position

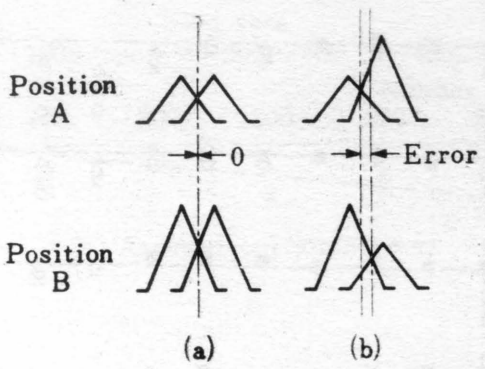


Fig 15 Error due to unbalance of output wave forms

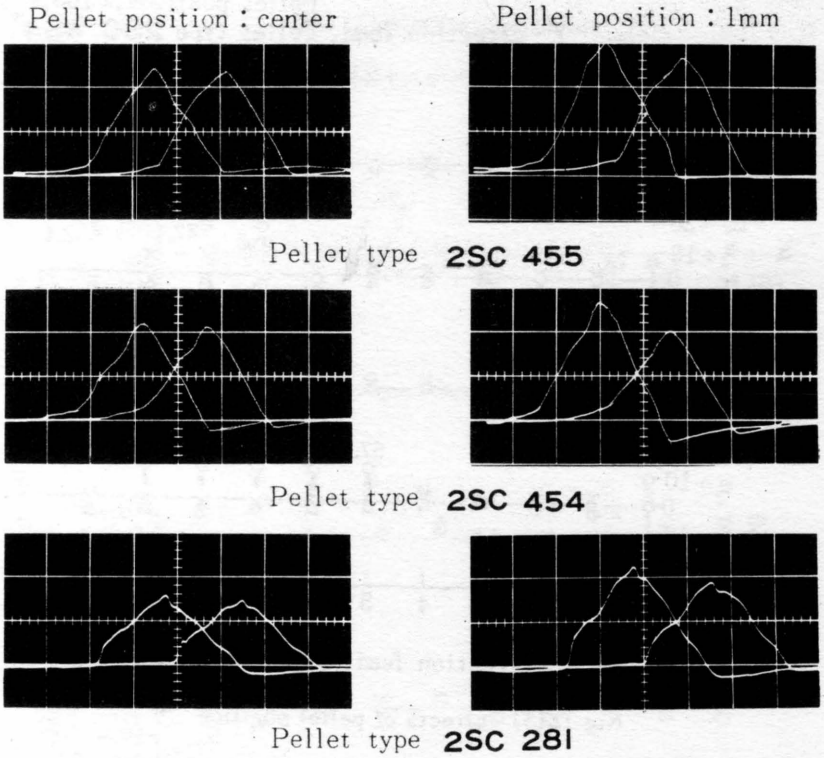


Fig 16 Output wave forms of photomultipliers

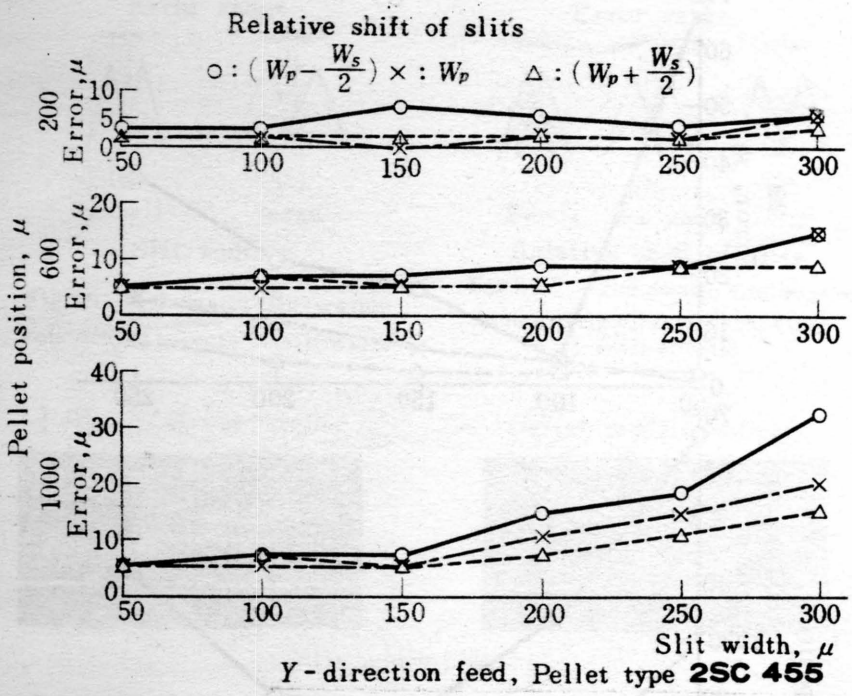


Fig 17(1) Effects of slit width and relative shift of slits

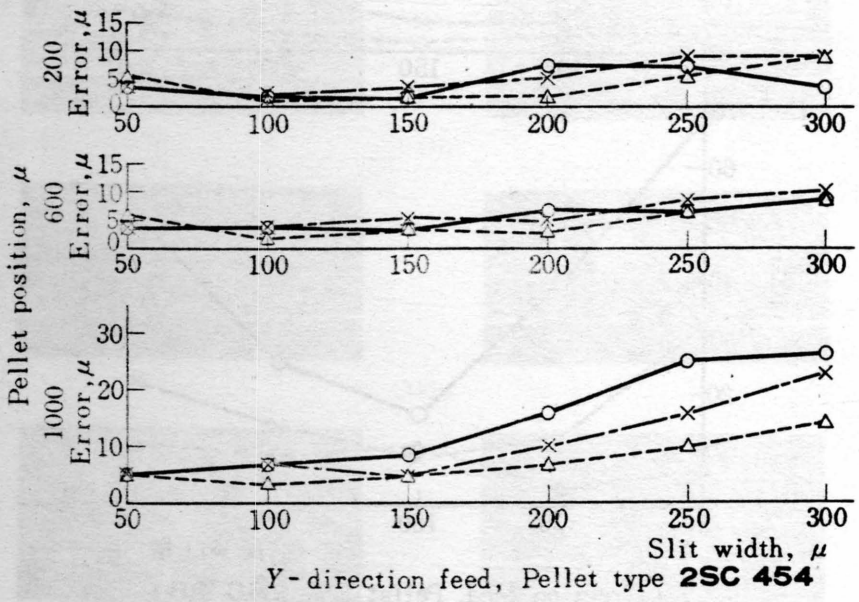
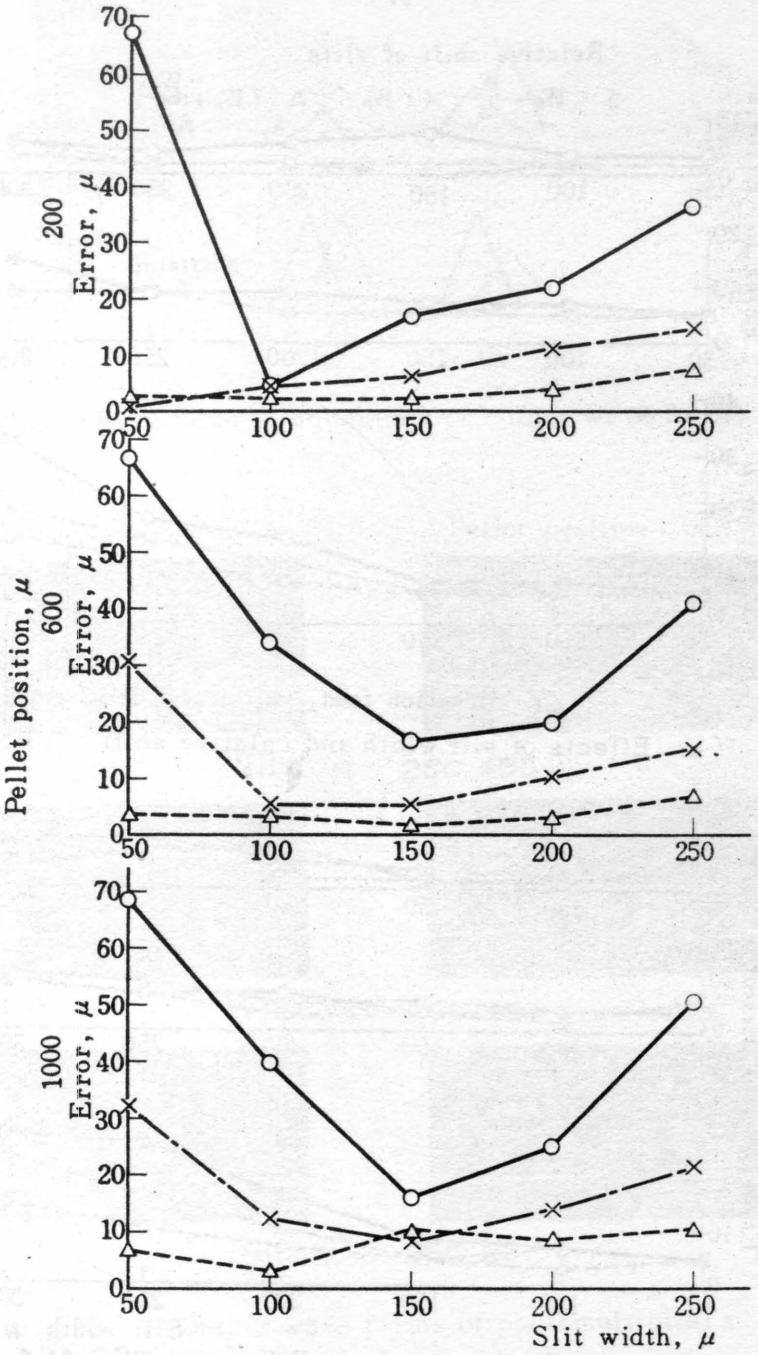


Fig 17(2)



Y-direction feed, Pellet type **2SC 281**

Fig 17(3)

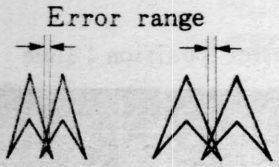


Fig 18 Schematic illustration concerning effects of slit width

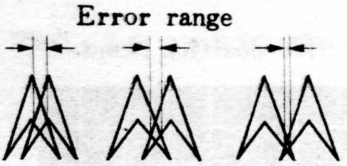


Fig 19 Schematic illustration concerning effects of relative shift of slits

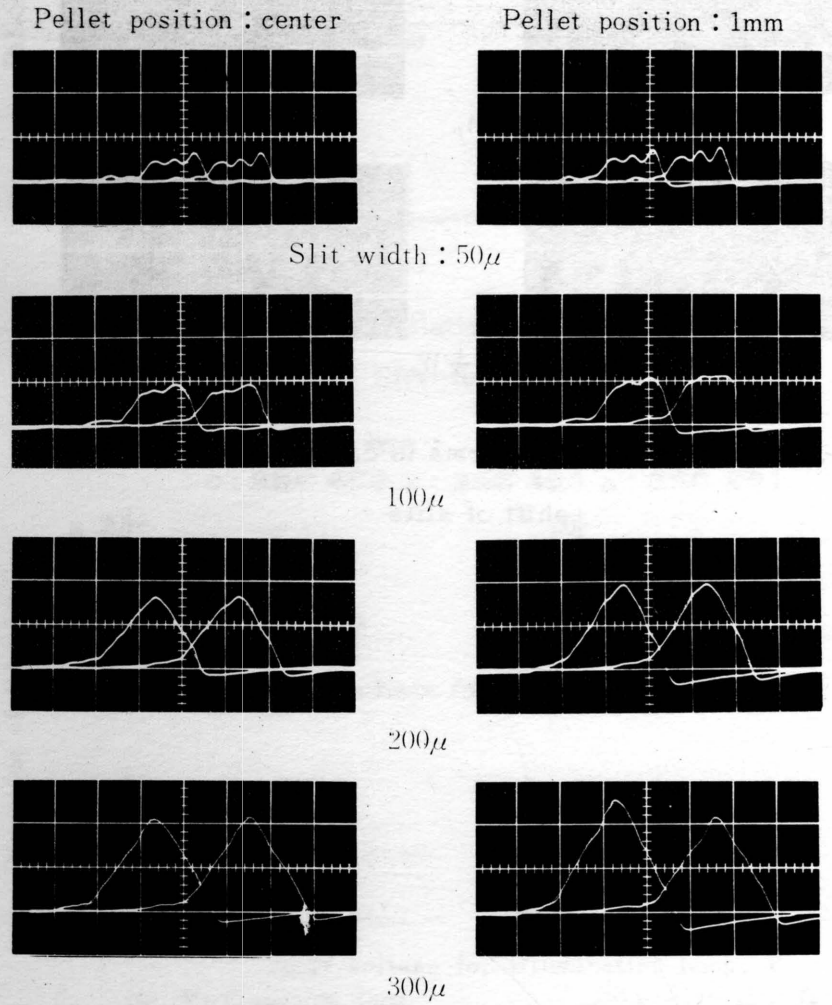


Fig 20 Output wave forms in cases of various slit widths

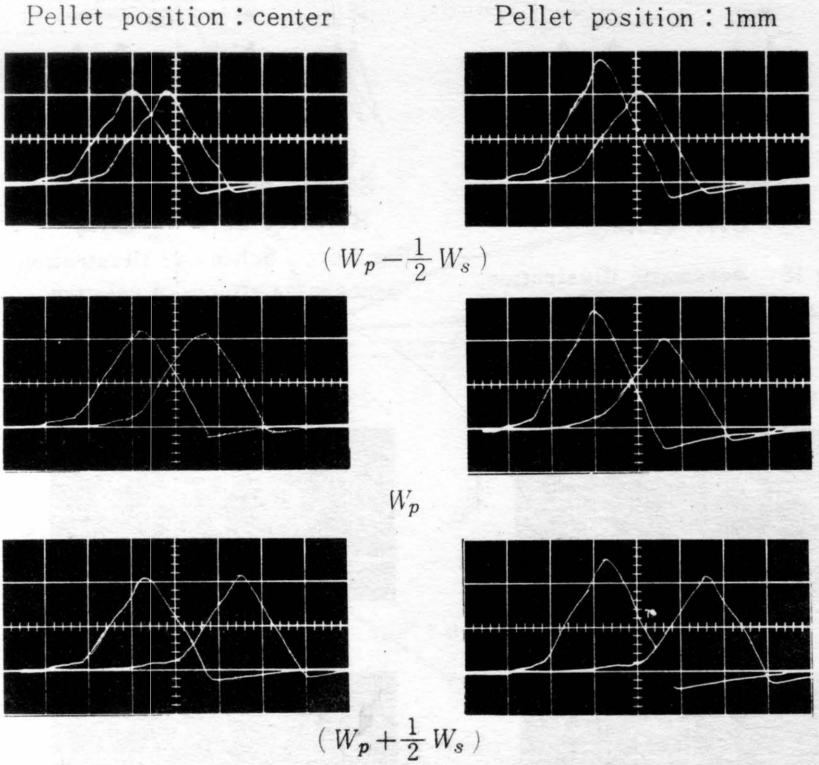


Fig 21 Output wave forms in case of each relative shift of slits

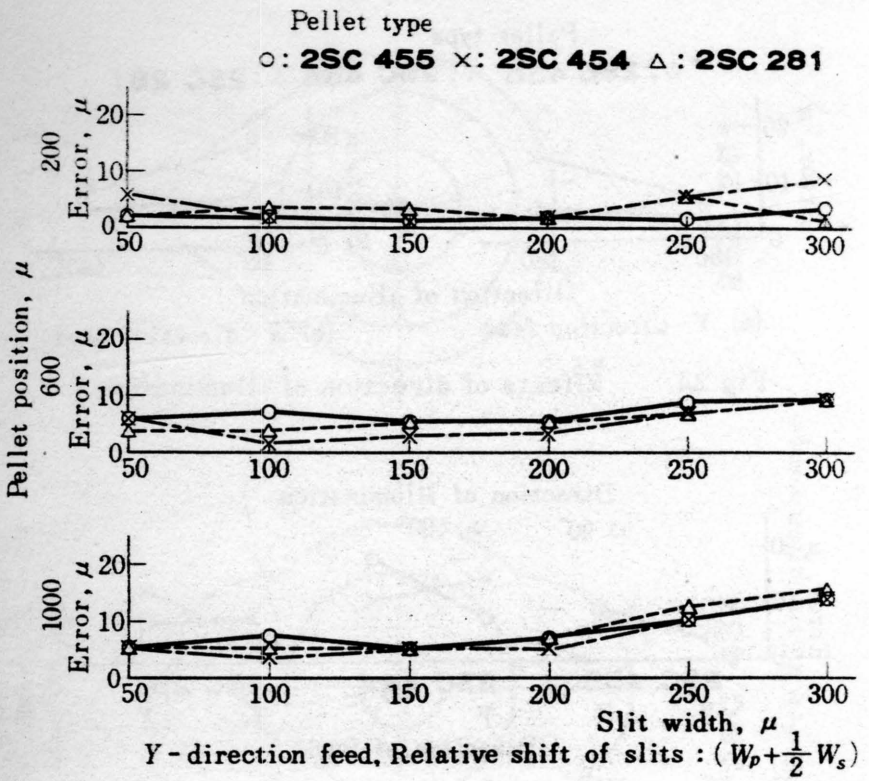


Fig 22 Effects of pellet type

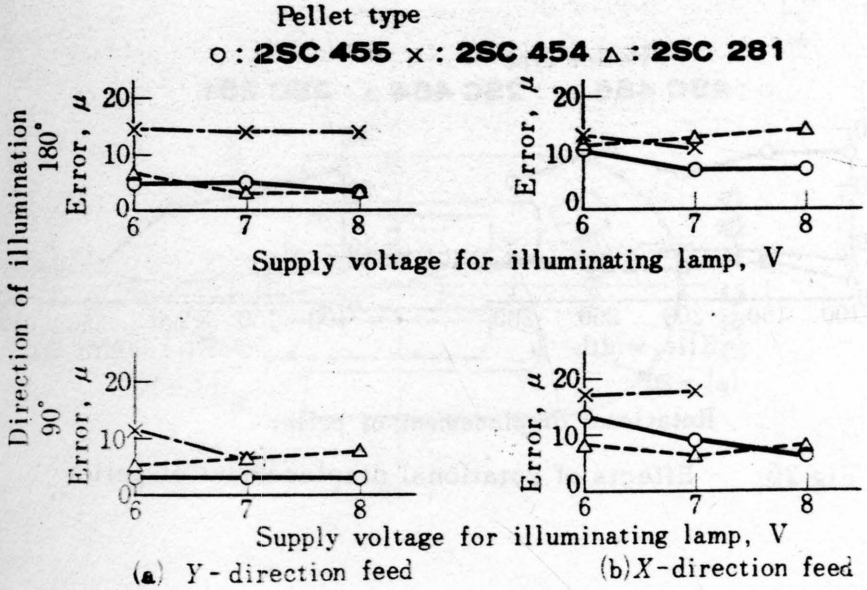


Fig 23 Effects of intensity of illumination

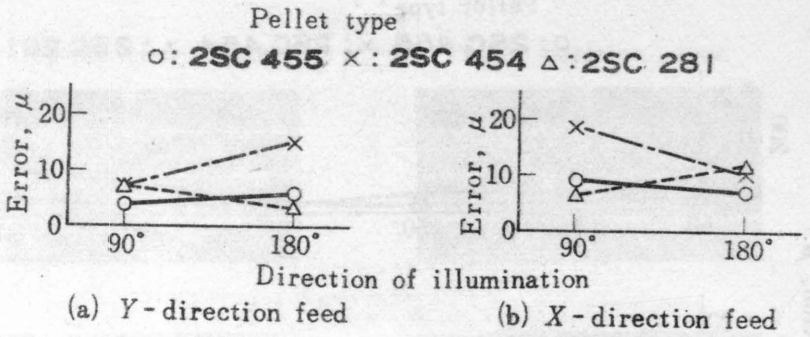


Fig 24 Effects of direction of illumination

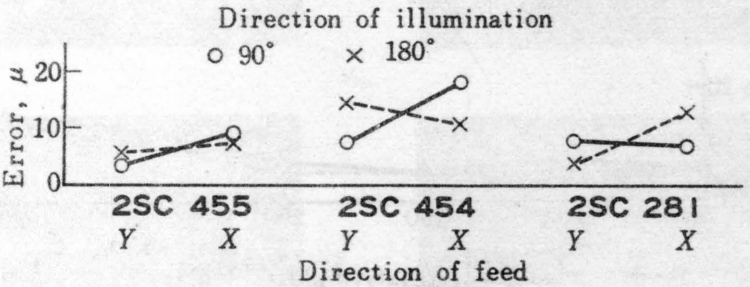


Fig 25 Effects of direction of illumination to each pellet type

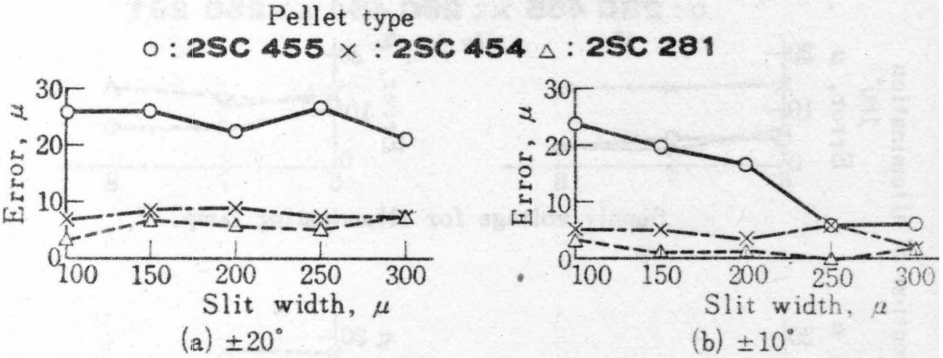


Fig 26 Effects of rotational displacement of pellet

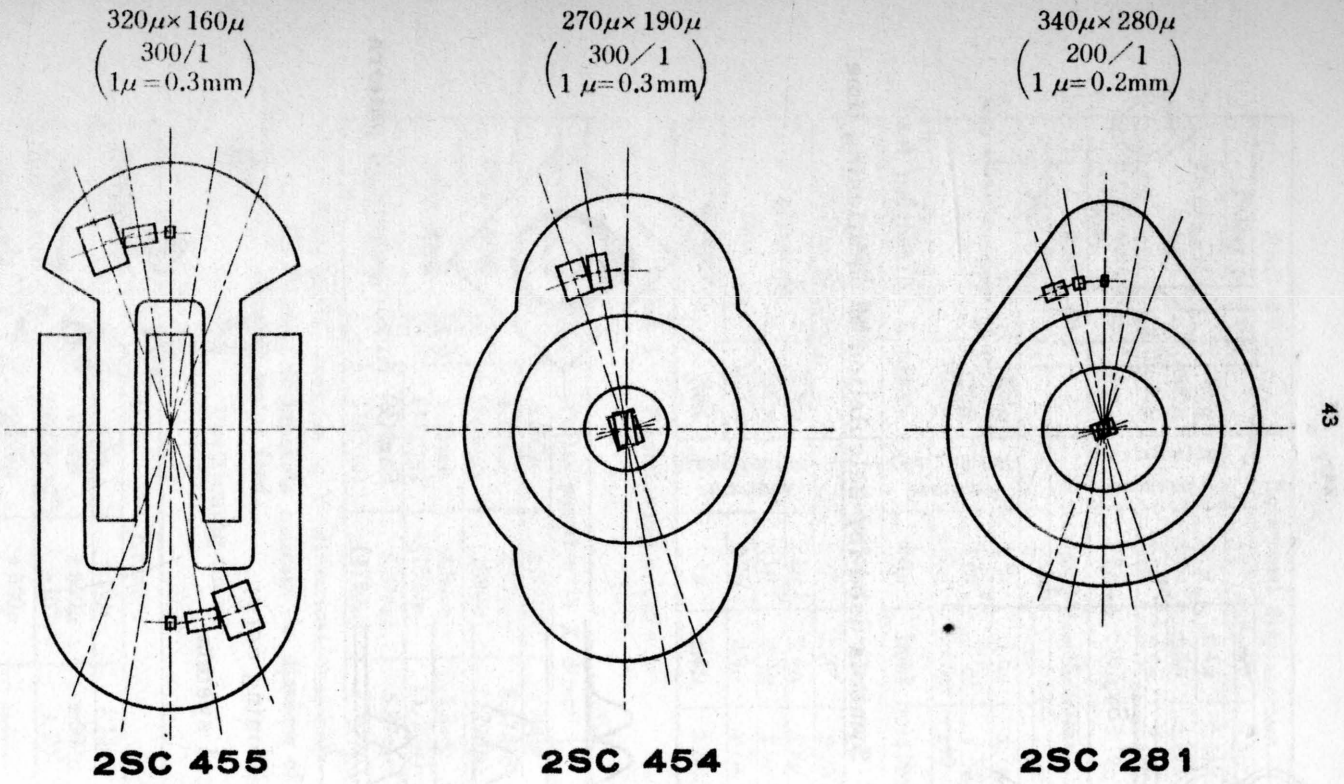
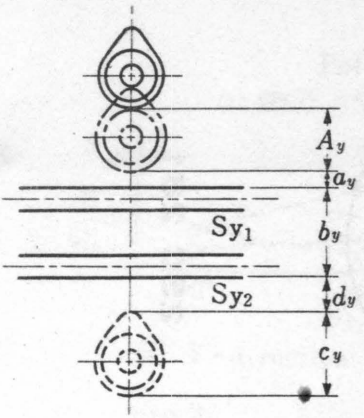
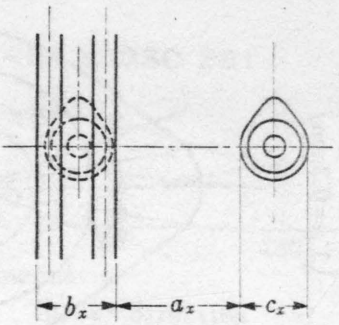


Fig 27 Error distribution range



(a) *Y*-direction feed



(b) *X*-direction feed

Fig 28 Symbols used for calculation of positioning time

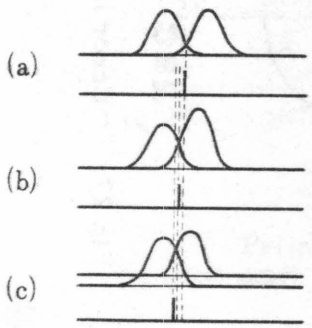


Fig 29 Positioning error from optical system

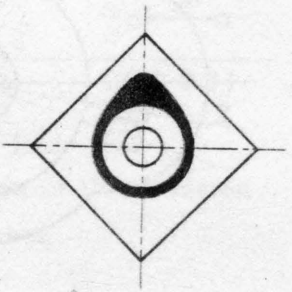


Fig 30 Arrangement of pattern

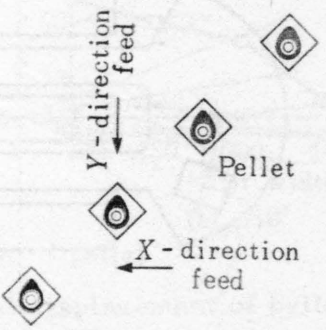


Fig 31
Arrangement of pellet in loading

2SC 455			Rotational displacement		
Pellet position			±0°	±10°	±20°
Y-direction feed	200μ	Positioning accuracy	±1μ	±4μ	±14μ
	600μ		±2μ	±5μ	±14μ
	1000μ		±2μ	±5μ	±15μ
X-direction feed	200μ		±2μ	±10μ	±12μ
2SC 454			Rotational displacement		
Pellet position			±0°	±10°	±20°
Y-direction feed	200μ	Positioning accuracy	±3μ	±5μ	±7μ
	600μ		±3μ	±6μ	±8μ
	1000μ		±8μ	±10μ	±10μ
X-direction feed	200μ		±0μ	±6μ	±8μ
2SC 281			Rotational displacement		
Pellet position			±0°	±10°	±20°
Y-direction feed	200μ	Positioning accuracy	±1μ	±4μ	±5μ
	600μ		±1μ	±4μ	±6μ
	1000μ		±2μ	±4μ	±7μ
X-direction feed	200μ		±1μ	±4μ	±10μ

Table 1 Positioning accuracy

Pellet type		Pellet position in X-direction		
		±200μ	±300μ	±500μ
2SC 455	Positioning time	4.72sec (1490μ)	4.97sec (1590μ)	5.47sec (1790μ)
2SC 454		4.55sec (1420μ)	4.8sec (1520μ)	5.3sec (1720μ)
2SC 281		5.12sec (1650μ)	5.37sec (1750μ)	5.87sec (1950μ)

※ (1) Initial position in Y-direction: 5mm±200μ
(2) Figures in brackets indicate the distance of stepping motor feed.

Table 2 Positioning time

Pellet type		Error		
		±5μ	±10μ	±15μ
2SC 455	Range of allowable pellet position	±500μ ±θ (0<θ<10°)	±500μ ±10°	±500μ ±20°
2SC 454		±500μ ±θ (0<θ<10°)	±500μ ±20°	——
2SC 281		±500μ ±10°	±500μ ±20°	——

Table 3 Range of allowable pellet position

INVESTIGATION OF A RECTIFIER REGULATING CIRCUIT AS A SAMPLED DATA SYSTEM

H. Bühler, Oerlikon Engineering Company, Zurich / Switzerland

1. General

1.1. Summary

Typical characteristics of rectifiers are discontinuous. If one device is ignited, then the process can only be altered again when the following device is ignited. Investigations into its transient conditions with the normal continuous methods, which use the mean value and modifications to voltage waveform, tend to give inexact results.

Corresponding to the descontinuous mode of the function of the rectifier, the circuit shows substantial similiarities to a sampled data system. However there are two basic differences from the normal linear sampled data system. On the one hand the sampling function is not composed of rectangular blocks, but of sinusoidal sections, and on the other, the sampling period is not constant when the ignition angle is altered.

The switching process of a rectifier is investigated as a sampled data system in reference [1], where it is shown that the parts of a sine wave can be absorbed relatively easily into the sampling function. The behaviour of a rectifier, as the ignition angle changes is determined in this paper. If the investigation is limited to small changes in the ignition angle, then it is possible, as will be shown later in the text, again using constant sampling periods, to determine the stability analysis for a rectifier regulating circuit just as for a normal sampling system.

In the following section 1.2. the well known voltage and current conditions for a rectifier are described and it is shown how the commutation can be approximated.

The behaviour of the rectifier is investigated in section 2 for small changes in firing angle, with special emphasis on the determination of the equivalent circuit diagram and the separate transfer functions.

Before the regulating circuit can be investigated in section 3, the

conditions in the gate control circuit must be determined (section 3.1.).

Finally, as an example, the results on a separately excited d.c. motor with rectifier supply and current control are quoted in section 3.5.

1.2. General voltage and current behaviour for a rectifier

1.2.1. Neglecting commutation

The following investigation is based on a p-phased controlled rectifier (see fig. 1 a). A back e.m.f. U_i is connected in series with a given load Z on the d.c. side. If it is assumed that only one individual device conducts and the commutation is neglected, then the short circuit inductance L_c on the d.c. side can be combined with the load Z to form the resulting impedance Z_r as is shown in fig. 1 b.

The rectifier voltage U_s, whose basic waveform for a two-phased controlled rectifier is shown in fig. 2 for distortion-free conduction, is superimposed onto the back e.m.f. U_i and the resulting impedance Z_r. The waveform of the direct current is also shown here schematically.

1.2.2. Approximate allowance for commutation

During continuous conduction two phases are short-circuited for short periods because of the commutation from one device onto the following one. Thus the rectifier voltage has the well-known waveform shown in fig. 3 a.

The rectifier voltage U_s can be considered as composed of two components; one component U_w which does not take the commutation into consideration (fig. 3 c) and a component U_k which allows for the additional voltage drop resulting from the commutation (fig. 3 d). These two components are built up of short sections of a sinusoidal function, whose widths correspond to the periods of commutation. Because the widths of this "commutation impulse" is normally not very large, it is arranged to replace the true waveform of finite width by a series of impulses of infinitely small width (Dirac Impulse). The area enclosed by these impulses is the same as the voltage-time area, for the voltage drop due to cummutation, which is dependent on the instantaneous value of the current at the instant of commutation. Thanks to this simplification the commutation can be allowed for with appreciably less trouble during the mathematical

treatment of the differential equations.

More exactly, during the period of commutation, the inductance L_c (see fig. 1 b) is less important on the d. c. side than the inductance $L_c/2$ [2] because of the two short circuited phases. However, to take the changes of the resulting load impedance into consideration would complicate the calculation very much so that only constant load impedance will be dealt with below (see fig. 1 b). The smallness of the error arising out of this will depend on the difference between the short circuit impedance L_c and the inductance of the load impedance Z.

2. Behaviour of the rectifier during small changes in the ignition angle

2.1. General

If the ignition angle remains constant, the rectifier operates like a sampled-data system with a sampling period of

$$T = \frac{2\pi}{p\,\omega_N} \tag{1}$$

where ω_N is the angular frequency of the network voltage.

However when the ignition angle changes, the sampling period is no longer constant and the problem cannot be solved in a general form. If the changes are limited to small values, as is usual in control technology, then the sampling period can again be regarded as constant.

To simplify the following, constant back e.m.f. U_i and continuous conduction are assumed. The changes in current then depend purely and solely on the change of ignition angle. The aim of this work is to find discrete transfer functions between the changes in ignition angle and current, whilst making allowance for commutation. Corresponding to the break down of the voltage U_s in fig. 3, into the components U_w (referred to in the following as the wave voltage) and the component U_k (referred to below as the commutation impulse), the change in current can also be built up from two components.

It should be mentioned that referred values are used. All the voltages are referred to a rated value U_n, and the currents to a rated value I_n.

2.2. Voltage changes

2.2.1. Changes in the wave voltage due to small changes in the ignition angle

We will determine first of all the changes in the wave voltage due to small changes in the ignition angle.

Fig. 4a shows a sinusoidal section from a sampling period for the wave voltage. The ignition angle changes by $\Delta\varphi_z$ and the extinction angle by $\Delta\varphi_1$, whereas the sinusoidal function is kept fixed relative to the time axes. This sets up a displaced sinusoidal section as in fig. 4b. The difference, i.e. the main voltage change Δu_w, is shown in fig. 4c. These small sinusoidal sections of Δu_w can be approximated, with a sufficiently good degree of accuracy, to Dirac-impulses with the same voltage-time area, for small changes in the ignition and extinction angles.

The voltage-time area referred to the sampling period T for an impulse due to a change in the ignition angle is

$$u_w \frac{\Delta\varphi_z}{\omega_N T} = - \frac{p}{2\pi} \hat{u} \sin \varphi_z \Delta\varphi_z \qquad (2)$$

and that for the "change in extinction angle"

$$u_w \frac{\Delta\varphi_1}{\omega_N T} = \frac{p}{2\pi} \hat{u} \sin\left(\varphi_z + \frac{2\pi}{p}\right) \Delta\varphi_1 \qquad (3)$$

If not only the individual parts of the sine wave are considered, but the whole waveform of the wave voltage, then it can be seen that at every sampling instant (for the same ignition point) an impulse for the change in ignition angle coincides with the change in extinction angle, for which $\Delta\varphi_1 = \Delta\varphi_z$ (see fig. 5a). Both of these impulses can be incorporated into a combined impulse with a voltage-time area of

$$u_w \frac{\Delta\varphi_z}{\omega_N T} = \frac{p}{\pi} \hat{u} \sin\frac{\pi}{p} \cos\left(\varphi_z + \frac{\pi}{p}\right) \Delta\varphi_z = \hat{u}_g \cos\left(\varphi_z + \frac{\pi}{p}\right) \Delta\varphi_z \qquad (4)$$

(see fig. 5b). Thus the maximum d.c. voltage obtained from the rectifier,

referred to the rated voltage U_n is combined with

$$\hat{u}_g = \frac{p}{\pi} \hat{u} \sin \frac{\pi}{p} \tag{5}$$

Changes in the ignition angle for wave voltages have the effect, as in a normal sampled-data system, of working with the Dirac impulses and the changes in the ignition angle are superimposed on the introduction of the impulses. The amplification factor of the sampling unit is thus

$$K_{go} = \hat{u}_g \cos(\varphi_z + \frac{\pi}{p}) \tag{6}$$

2.2.2. Changes in voltage due to the commutation impulses

As is shown in fig. 3e, the commutation can be considered, to a good degree of approximation, as a Dirac impulse. The area swept out by this "commutation impulse" is equal to the voltage-time area of the inductive voltage drop. The voltage-time area referred to the sampling period is equal to $r_i \cdot i$, where

$$r_i = \frac{p}{2\pi} \omega_N L_c \frac{I_n}{U_n} \tag{7}$$

which is the equivalent internal resistance of the rectifier expressed in referred values, and i the commutating current.

Thus the effect of commutation can thereby be represented by a normal sampling system with Dirac impulses, where the amplification factor for the sampling unit is equal to r_i and onto which the current i is superimposed.

Special attention must be paid to the determination of the determinate current change for the commutation during small changes in the ignition angle. The conditions should be investigated with the help of fig. 6. This shows the behaviour of the current with respect to time. The full line shows the steady state behaviour before the ignition angle is changed and the broken line shows its behaviour after the change in ignition angle, designated by i. The difference $\Delta i = i - i_{stat}$ is shown by a dot-dash line. The change of ignition angle during the sampling period is

$$\Delta \varepsilon_z = \frac{p}{2\pi} \Delta\varphi_z \tag{8}$$

The determinate current for the commutation in the steady state condition is $i_{stat}[n,0]$ whereas during a change in the ignition angle the current $i[n,\Delta\varepsilon_z]$ must be considered. The effective change in current for the commutation is therefore

$$\Delta i_c[n] = i[n,\Delta\varepsilon_z] - i_{stat}[n,0] \tag{9}$$

As can be seen from fig. 6, this current change can also be written

$$\Delta i_c[n] = \Delta i[n-1,\,1] + \frac{\partial i[n-1,\,1]}{\partial \varepsilon}\,\Delta\varepsilon_z[n] \tag{10}$$

Because small changes were specified, the current i differs by only a small amount from the steady state current i_{stat} and the gradient $\partial i/\partial\varepsilon$ can be replaced by $\partial i_{stat}/\partial\varepsilon$. If the change in ignition angle $\Delta\varepsilon_z[n]$ is replaced by $\Delta\varphi_z[n]$, then it follows that

$$\Delta i_c[n] = \Delta i[n-1,\,1] + K_{cz}\,\Delta\varphi_z[n] \tag{11}$$

where

$$K_{cz} = \frac{p}{2\pi}\,\frac{\partial i_{stat}[n-1,\,1]}{\partial\varepsilon} \tag{12}$$

2.3. The block diagram

As was shown in the previous section, the behaviour of the rectifier during small changes in the ignition angle can be represented by two sampling units, one of which emits pulses as a result of changes in the ignition angle, and the other emits impulses because of commutation. The latter behave exactly as in equation (11) and are directly dependent on changes in the current as well as the ignition angle. A block diagram can be set up, as shown in fig. 7 a. G_i is the transfer function for the load in the d.c. circuit, i.e. the impedance Z_r in fig. 1 b.

Because both of the sampling units operate synchronously and in phase, the subtraction point in front of the sampling units can be transferred, which means that they can be combined into a single sampling unit. Both parallel branches, which are directly dependent on the change in firing angle can be united into a block with an amplification factor of

$$K_g = K_{go} - r_i \cdot K_{cz} \tag{13}$$

The block diagram, transformed and simplified as above, is illustrated in fig. 7b.

2.4. The discrete transfer function

The simplified block diagram for the rectifier represents a feedback sampling system. Further investigations can therefore be made using the usual methods for sampling systems, in particular with the aid of a discrete Laplace transformation (see [3] and [4]).

It should be noted when setting up the system of equations, that because of the Dirac impulses, the change in current Δi can produce a spike just at the instant of sampling. According to equation (11) the current change should be obtained just before the instant of sampling, and the change in ignition angle immediately afterwards. Using the relationship between the original function and the discrete function in the q-plane

$$\left.\begin{aligned} \Delta\varphi_z [n] \circ\!\!-\!\!\bullet\ \Delta\varphi_z^* (q) \\ \Delta i [n-1, 1] \circ\!\!-\!\!\bullet\ \Delta i^* (q,1)\, e^{-q} \end{aligned}\right\} \tag{14}$$

the change in current $\Delta i^*(q,\varepsilon)$ in the q-plane of the discrete Laplace transformation follows from

$$\Delta i^* (q,\varepsilon) = K_i^* (q,\varepsilon) \left[K_g \Delta\varphi_z^* (q) - r_i \Delta i^* (q,1)\, e^{-q} \right] \tag{15}$$

where $K_i^*(q,\varepsilon)$ is the discrete transfer function for the d.c. load. Should the normal transfer function be a rational function of the form

$$G_i (q) = \frac{q_i (q)}{p_i (q)} \tag{16}$$

then the discrete transfer function follows from

$$K_i^* (q,\varepsilon) = \sum_{k=1}^{l} \frac{q_i (q_k)}{p_i' (q_k)} \frac{e^q}{e^q - e^{q_k}} e^{q_k \varepsilon} \tag{17}$$

q_k are the l zero points from $p_i(q)=0$. Thus the summation covers all of the l zero points. It is assumed in equation (17) that no multiple zero points occur (if so, see [3] and [4]).

Equation (15) can be transformed by letting $\varepsilon = 1$. From this $\Delta i^*(q,1)$

can be found. If this intermediate result is substituted into equation (15), then the discrete transfer function for small changes in the firing angle can be obtained very quickly

$$K_{\Delta}^{*}(q,\varepsilon) = \frac{\Delta i^{*}(q,\varepsilon)}{\Delta\varphi_z^{*}(q)} = \frac{K_g K_i^{*}(q,\varepsilon)}{1 + r_i e^{-q} K_i^{*}(q,1)} \tag{18}$$

2.5. The transfer function

The transfer function represents the behaviour of the change in current Δi with respect to time, referred to the change in the ignition angle, when the ignition angle changes $\Delta\varphi_z$ in the form of a step function. The transfer function follows from the discrete transfer function by back-transformation

$$\frac{\Delta i[n,\varepsilon]}{\Delta\varphi_z} \quad o\!-\!\!-\!o \quad K_{\Delta}^{*}(q,\varepsilon) \quad \frac{e^{q}}{e^{q}-1} \tag{19}$$

In the case of an ohmic-inductive d.c. load with a time constant τ_o referred to the sampling period T, which is obtained from the normal transfer function

$$G_i(q) = \frac{1}{r} \quad \frac{1}{1 + q\tau_o} \tag{20}$$

the transfer $G_i(q)$ function can be determined. The behaviour of the change in the current factor

$$\Delta k_i[n,\varepsilon] = \frac{r\,\Delta i[n,\varepsilon]}{\hat{u}} \tag{21}$$

with respect to time referred to the change in the ignition angle $\Delta\varphi_z$ is shown in fig. 8, and in particular the effect of the commutation is investigated. It can be clearly seen that the change in current decreases very quickly as the ratio $r_i/r\tau_o$ is increased. For a steady time constant τ_o, the transient process reduces more quickly the larger $r_i/r\tau_o$ becomes. When $r_i/r\tau_o = 1$, the final stationary process will be reached right from the start. $r_i/r\tau_o$ implies that the d.c. load is purely ohmic and that only the short circuit inductance L_c is evident.

3. The control circuit

3.1. Conditions in the gate control circuit for small control voltage changes

Before the investigation into the control circuit can be started,

special attention must be paid to the gate control circuit. An ignition pulse is conducted to the rectifier connected next in circuit, from a continuously changing control voltage u_s in the gate control circuit, where the level of the ignition impulse, i.e. the ignition angle φ_z, is a function of the control voltage u_s. Practically no time constants occur in a transistorized gate control circuit. When determining the transfer function, one must appreciate that the control voltage u_s normally has a ripple. The corresponding conditions should be found with the help of fig. 9.

The behaviour of the control voltage for the steady state condition $u_{s\,stat}$ is shown, and the dashed line represents a small change (u_s). The difference Δu_s is shown by a dot-dash line. The intersection of the control voltage with a saw-tooth shaped reference voltage u_r, for example, gives the level of the impulses. The change in the control voltage important for the change in the ignition angle (expressed by $\Delta\varepsilon_z$) follows according to fig. 9 from

$$\Delta u_{sz}[n] = \Delta u_s[n-1,\,1] + \frac{\partial u_s[n-1,\,1]}{\partial \varepsilon}\,\Delta\varepsilon_z[n] \qquad (22)$$

If small changes are again considered, then u_s does not differ greatly from $u_{s\,stat}$ and $\partial u_s/\partial\varepsilon$ can be replaced by $\partial u_{s\,stat}/\partial\varepsilon$. On the other hand the relation

$$\Delta\varphi_z[n] = -K_s\,\Delta u_{sz}[n] \qquad (23)$$

is obtained from the gradient of the reference voltage u_r. Taking equation (8) into consideration, the resulting transfer function for the gate control circuit is obtained from

$$K_{sr} = \frac{\Delta\varphi_z[n]}{\Delta u_s[n-1,1]} = -\frac{K_s}{1+\frac{p}{2\overline{u}}K_s\frac{\partial u_{sstat}[n-1,1]}{\partial\varepsilon}} \qquad (24)$$

The block diagram shown in fig. 10 can be used for the gate control circuit. The fact that the gate control circuit operates like a sampling circuit will be taken into consideration by a sampling unit connected immediately in circuit with Dirac impulses.

3.2. The block diagram for the control circuit

By using diagrams 7b and 10, the block diagrams shown in fig. 11a

for the control circuit, can be set up. The regulator has a transfer function G_R, and smoothing circuits with the transfer function G_f are provided in the feedback loop. The block diagram can be transformed and simplified.

Both sampling units work synchronously and in phase. Because a unit with pure P-conditions (transfer function K_g) lies between the two sampling units, the first unit can be removed. Both blocks with the transfer factors K_{sr} and K_g can be combined into a single unit with the transfer factor

$$K = K_{sr} \cdot K_g \tag{25}$$

The junction at the output, i.e. behind the block G_i, can be put in front of this, which means that both the feedback circuits must each have a unit with the transfer function G_i inserted into them. From this the block diagram in fig. 11 b follows.

A further step is to transfer the subtraction point for the reference-actual value comparison unit by two blocks to the right. These two blocks must also be added to the lower feedback circuit (fig. 11 c).

Both parallel feedback loops can be combined into a single block, having the transfer function

$$G(q) = K\, G_R(q) \cdot G_f(q) \cdot G_i(q) + r_i\, G_i(q) \tag{26}$$

However both the blocks at the inlet to the reference value change Δi_c can be united into a single unit with the transfer function

$$G_c(q) = K \cdot G_R(q) \tag{27}$$

This is how the completely reduced block diagram shown in fig. 11 d is obtained.

3.3. The function in the q-plane for the change in current

Before the function in the q-plane for the change in current can be found, the relationship between the change in the reference value Δi_c and the sampling points (ignition points for the rectifiers) with respect to time, must be considered.

The behaviour of the change in reference value Δi_c with respect to

time is drawn schematically in fig. 12. The zeropoint of the reference time $t/T = n + \varepsilon$ is synchronized with the start of the reference value change. Δi_c is transformed into the change Δx_c by the unit G_c in the block diagram in fig. 11 d. $\Delta x_\varepsilon = \Delta x_c - \Delta x$ is sampled by means of the sampling unit. Normally the start of the change in the reference value is displaced from the instant of sampling by ε_c. This displacement can, within equal statistical probability, have values in the range $0 \leq \varepsilon_c \leq 1$. ε_c can have a mean value of 1/2. The sampling does not take place with integer values of n, but with $n + \varepsilon_c$ (see fig. 12).

Corresponding to the arrangement in fig. 11 d, the following relationships can be set up for the block region in discrete Laplace transformations

$$\Delta i^*(q,\varepsilon) = K_i^*(q, \varepsilon - \varepsilon_c)\, \Delta x_\varepsilon^*(q, \varepsilon_c) \tag{28a}$$

$$\Delta x_\varepsilon^*(q, \varepsilon_c) = \Delta x_c^*(q, \varepsilon_c) - \Delta x^*(q, \varepsilon_c) \tag{28b}$$

$$\Delta x^*(q,\varepsilon) = K^*(q, \varepsilon - \varepsilon_c)\, \Delta x_\varepsilon^*(q, \varepsilon_c) \tag{28c}$$

$K_i^*(q,\varepsilon)$ and $K^*(q,\varepsilon)$ are the discrete transfer functions relating to $G_i(q)$ and $G(q)$ respectively [see also equation (17)]. According to [4]

$$K^*(q, \varepsilon - \varepsilon_c) = e^{-q} K^*(q, 1 + \varepsilon - \varepsilon_c) \qquad 0 \leq \varepsilon \leq \varepsilon_c \tag{29a}$$

$$= K^*(q, \varepsilon - \varepsilon_c) \qquad \varepsilon_c < \varepsilon \leq 1 \tag{29b}$$

In order to determine the value $\Delta x^*(q, \varepsilon_c)$ at the instant of sampling, it must be noted that the value corresponding to the change in current must be obtained immediately before the sample (see section 2.4.). The relationship (29a) must be used, where ε must be replaced by ε_c. Thus it follows from equation (28c)

$$\Delta x^*(q, \varepsilon_c) = e^{-q} K^*(q,1)\, \Delta x_\varepsilon^*(q, \varepsilon_c) \tag{30}$$

Substituting this relationship into equation (28b) and solving for $\Delta x_\varepsilon^*(q, \varepsilon_c)$ and using the answer in (28a), the change in current becomes

$$\Delta i^*(q,\varepsilon) = \frac{K_i^*(q, \varepsilon - \varepsilon_c)}{1 + e^{-q} K^*(q,1)}\, \Delta x_c^*(q, \varepsilon_c) \tag{31}$$

For a given relationship concerning the change in the reference

value Δi_c with respect to time, $\Delta x_c^*(q, \varepsilon_c)$ can be found. In particular, for a step function change in the reference value Δi_{co} and a PI regulator having the normal transfer function

$$G_R(q) = \frac{1}{q\tau_i} + \frac{\tau_1}{\tau_i} \tag{32}$$

the following relationship applies

$$\Delta x_c^*(q, \varepsilon_c) = K\left[\frac{1}{\tau_i}\frac{1}{e^q-1} + \frac{\tau_1+\varepsilon_c}{\tau_i}\right]\frac{e^q}{e^q-1}\Delta i_{co} \tag{33}$$

The dead time ε_c has the same effect as an increase in displacement period τ_1.

For the stability of current control, the characteristic equation

$$1 + e^{-q} K^*(q,1) = 0 \tag{34}$$

holds, according to equation (31). The usual methods for sampled-data systems may be used for the stability investigation. Refer to [4], in this connection.

The behaviour of the change in current relative to time can be obtained from equation (31) by the use of the inverse discrete Laplace transformation (see [4]). This is relatively easily obtained for a definite case, i.e. for given values of $K_i^*(q,\varepsilon)$ and $K^*(q,\varepsilon)$ of not too high an order.

3.4. Back transformation in the q-plane of normal Laplace transforms

In many ways the control circuit is similar to the intermost control loop for a cascade regulator, upon which further regulating circuits are superimposed. The superimposed regulating circuits operate continually, and energy sources are to be found between the regulated current and the superimposed regulated objects, which practically smooth out all the effect of ripple on the current. Therefore it is desirable that the superimposed regulating circuits should be investigated using the normal methods valid for continuous, linear regulating circuits, especially for the frequency-response characteristic. For this it is necessary to transform the transfer function for the closed regulating circuit of the q-plane for the discrete Laplace transform back within the q-plane for the usual Laplace transform. According to [4] the D^{-1} transform should be used

for this. Therefore

$$\tilde{f}(q) = \int_0^1 e^{-q\varepsilon} F^*(q,\varepsilon)\, d\varepsilon \qquad (35)$$

Inserting the expression for $\Delta i^*(q,\varepsilon)$ in equation (31) into $F^*(q,\varepsilon)$ and remembering that only $K_i^*(q,\varepsilon-\varepsilon_c)$ dependent on ε, then

$$\Delta\tilde{i}(q) = \frac{\Delta x_c^*(q,\varepsilon_c)}{1+e^{-q}K^*(q,1)} \int_0^1 e^{-q\varepsilon} K_i^*(q,\varepsilon-\varepsilon_c)\, d\varepsilon \qquad (36)$$

Care must be taken in the integration that either $e^{-q}K_i^*(q,1+\varepsilon-\varepsilon_c)$ or $K_i^*(q,\varepsilon-\varepsilon_c)$ should be used, according to the limits in which ε lies, corresponding to equation (29) for $K_i^*(q,\varepsilon-\varepsilon_c)$. If $K_i^*(q,\varepsilon)$ is replaced by the summation in equation (17), then by integrating further

$$\int_0^1 e^{-q\varepsilon} K_i^*(q,\varepsilon-\varepsilon_c)\, d\varepsilon = \sum_{k=1}^{l} \frac{q_i(q_k)}{p_i'(q_k)} \frac{e^{-q\varepsilon_c}}{q-q_k} = G_i(q)\, e^{-q\varepsilon_c} \qquad (37)$$

from this the transfer function is obtained

$$G_i(q) = \frac{q_i(q)}{p_i(q)} = \sum_{k=1}^{l} \frac{q_i(q_k)}{p_i'(q_k)} \frac{1}{q-q_k}. \qquad (38)$$

The factor $e^{-q\varepsilon_c}$ in equation (37) allows for the dead time between the change in the reference value and the sampling point.

The change in current $\Delta J(q)$ is obtained by inserting the relationship (37) into equation (36), in the q-plane of the usual Laplace transform. It should be noted that both before and afterwards that $\Delta x^*(q,\varepsilon_c)$ and the denominator $1 + e^{-q}K^*(q,1)$ are functions of e^q.

However it is desirable to express the transfer function for the closed control circuit as a rational function of q. This can be approximated. The most obvious approximation is by putting $e^q \approx 1 + q$. However it turns out that this approximation is not very good, and yields results which are unsatisfactory in practice. A very good approximation can be obtained, however, using a function of the form

$$e^q \approx \frac{2+q}{2-q} \qquad (39)$$

for abbriviation we insert

$$e^{-q} K^*(q,1) = K_s^*(q,1) = K_s^*(e^q) \approx G_s^*\left(\frac{2+q}{2-q}\right) = G_s^*(q) \tag{40}$$

The star in $G_s^*(q)$ is for remembering that this transfer function was obtained from the discrete transfer function $K_s^*(q,1)$.

If (39) is put into (33) then it follows quickly:

$$\Delta x_c^*(q, \varepsilon_c) = K\left[\frac{1}{q\tau_i} + \frac{\tau_1 + \varepsilon_c - \frac{1}{2}}{\tau_i}\right](1 + \frac{q}{2})\ \frac{\Delta i_{co}}{q} \tag{41}$$

Using the mean value of $\varepsilon_c = 1/2$, and noting that $\Delta i_{co} q = \Delta \tilde{i}_c$, it follows that:

$$\Delta x_c^*(q, \varepsilon_c) = K\left(\frac{1}{q\tau_i} + \frac{\tau_1}{\tau_i}\right)(1 + \frac{q}{2})\ \Delta\tilde{x}_c = K\, G_R(q)\,(1 + \frac{q}{2})\ \Delta\tilde{x}_c \tag{42}$$

It has been found that this relation also applies for changes in the reference value other than step functions, within a good degree of accuracy.

Finally the normal transfer function for the closed loop control circuit follows from

$$\frac{\Delta\tilde{x}(q)}{\Delta\tilde{x}_c(q)} = \frac{K\, G_R(q) \cdot G_i(q)(1 + \frac{q}{2})}{1 + G_s^*(q)}\ e^{-\frac{q}{2}} \tag{43}$$

This keeps the effect of the mean dead time $\varepsilon_c = 1/2$ in exponential form $e^{-q/2}$.

As opposed to the normal continual treatment of the control circuit, the additional factor $(1 + q/2)$ appears in the numerator, and the denominator has $1+G_s^*(q)$ instead of $1+KG_RG_ie^{-q/2}$, where $G_s^*(q)$ comes from the discrete Laplace transform [see equation (40)]. $G_s^*(q)$ also covers supplementary the effect of commutation.

3.5. Example: Control circuit for a separately excited d. c. motor using a rectifier supply

Finally an important application is shown as an example, namely the control circuit for a separately excited direct current motor with a rectifier supply.

Taking any possible change of speed into consideration, the transfer function for the d.c. load appears as (see [5])

$$G_i(q) = \frac{q\tau_m}{1 + q\, r_a \tau_m (1 + q\tau_a)} \qquad (44)$$

r_a is the ohmic resistance of the armature circuit, $\tau_a = T_a/T$ is the referred armature time constant, and $\tau_m = T_m/T$ the mechanical time constant. A PI regulator is need for regulation having the transfer function in equation (32).

It is beyond the scope of this paper to write the transfer function out in detail. Fig. 13 only shows the unit-step response obtained from back transforming the transfer function for a step change in the reference value.

Three curves are drawn in fig. 13. Curve 1 comes from back transforming the discrete function in the q-plane in equation (31). It can be seen that the ripple increases with time. This is also obvious from physical considerations. If the opposing torque is independent of speed, the speed of the motor increases linearly when the armature current is increased uniformly. However because the change in current remains constant, the ignition angle must be altered continuously, which has the result that the ripple due to the current change increases all the time.

Curve 2 in fig. 13 is found by back transforming the normal transfer function in equation (43) for a step change in the reference value. It can be seen that the actual mean shape is a very good approximation. Other examples have also shown that back transformation in the q-plane for normal Laplace transforms as dealt with in section 3.4., together with the approximation $e^q = (2+q)/(2-q)$ yields very good results for practical applications.

Finally curve 3 in fig. 13 shows the unit-step response again, which is obtained from the normal static treatment of the rectifier control circuit (see [5]). The control circuit is treated in this as a continuous system, and also the commutation is not considered. It can be seen that as opposed to the true shapes (curves 1 and 2) a definite oscillation occurs for the same regulator settings. The static treatment by itself of rectifier control circuits only gives very approximate results which worsen, the

larger the sampling period T is in comparison on the time constants in the circuit.

Literature

[1] BAYOUMI, M.: Transient and steady state current of a q-phase controlled rectifier. Neue Technik (1962), H. 3, S. 141 - 145.

[2] WASSERRAB, TH.: Schaltungslehre der Stromrichtertechnik. Springer, Berlin/Göttingen/Heidelberg, 1962

[3] ZYPKIN, J.S.: Differenzgleichungen der Impuls- und Regeltechnik. VEB-Verlag, Berlin (1956)

[4] ZYPKIN, J.S.: Theorie der linearen Impulssysteme. Oldenburg, München (1967).

[5] BUEHLER, H.: Einführung in die Theorie geregelter Gleichstromantriebe. Birkhäuser, Basel (1962)

Subscripts on the Figures

Fig. 1. p-phased controlled rectifier.
a) connection diagram.
b) combination of the short circuit inductance L_c into the resulting impedance $Z_r = Z + sL_c$ on the d.c. side.

Fig. 2. Voltage and current curves neglecting commutation and having continuous conduction.
a) Rectifier voltage U_s.
b) D.C. current I.

Fig. 3. Approximate allowance for commutation. (Separation into wave voltage and commutation impulse).
a) Rectifier voltage U_s.
b) D.C. current I.

c) Wave voltage U_w.

d) Components of the voltage drop because of commutation U_k.

e) Replacing the sinusoidal section of U_k by Dirac impulses.

Fig. 4. Voltage change in the wave voltage U_w as a result of changes in the ignition and extinction angles $\Delta\varphi_Z$ and $\Delta\varphi_L$ respectively

a) Sinusoidal section for the original conditions.

b) Sinusoidal section for a change in ignition or extinction angle.

c) Voltage difference Δu_w.

d) Replacing by smoothing Dirac impulses.

Fig. 5. Superimposing the Dirac impulses for changes in the ignition and extinction angles (a) onto a resulting impulse (b).

Fig. 6. Determination of the main current change for the commutation $\Delta i_c[n]$.

Fig. 7. Block diagram for the rectifier, for small changes in ignition angle.

a) Detailed block diagram.

b) Simplified block diagram.

Fig. 8. Transfer function for the change in the current factor $\Delta k_i[n, \varepsilon]$, referred to the change in ignition angle $\Delta\varphi_Z$. Ohmic-inductive d.c. load, time constant $T_0 = 1,5$ referred to the sampling period. Effect of the internal resistance r_i (commutation).

Fig. 9. Determination of the transfer function for the gate control circuit.

Fig. 10. Block diagram for the gate control circuit.

Fig. 11. Block diagram for the control circuit.

a) Detailed block diagram.

b) and c) Intermediate steps in the simplification (see text).

d) Completely reduced block diagram.

Fig. 12. Relationship between sampling point and change in reference value.

Fig. 13. Speed regulation for a.d.c. motor with the control circuit data

as given, equalizing process for the closed loop with a step change in the reference value.

1) Time behaviour of the sampling system.
2) From the function in the q-plane for the normal Laplace transform [equation (43)] .
3) From the usual continuous treatment of the rectifier control circuit.

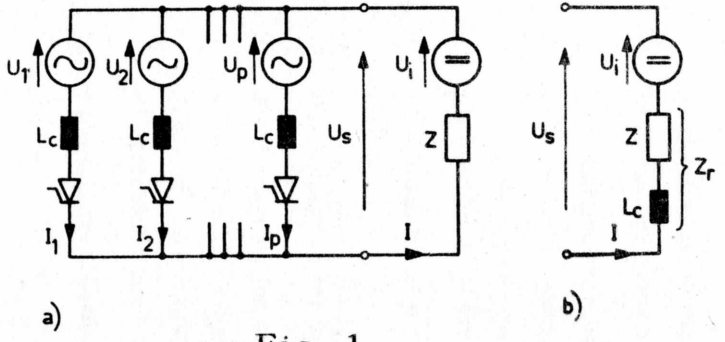


Fig 1

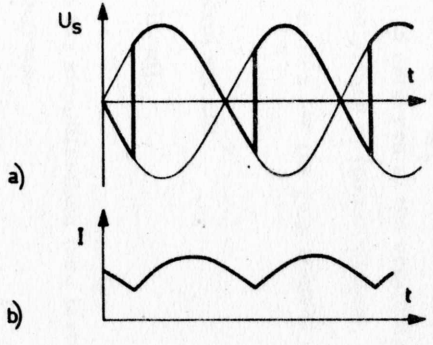


Fig 2

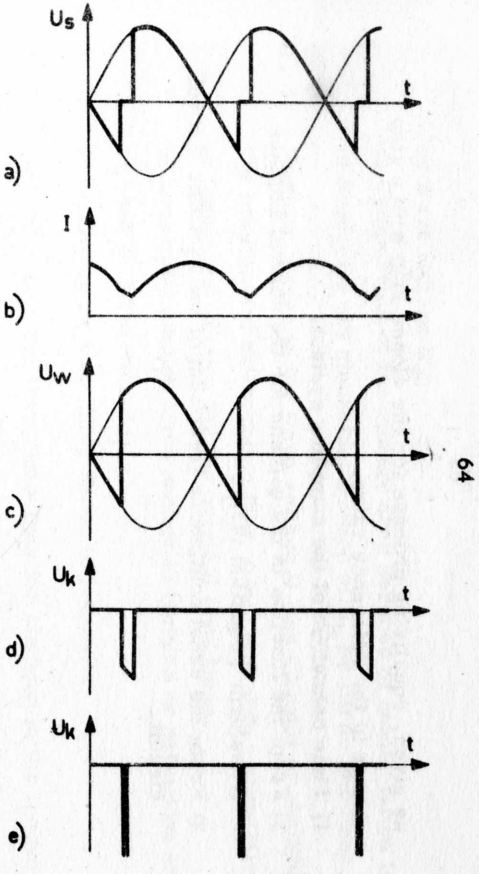


Fig 3

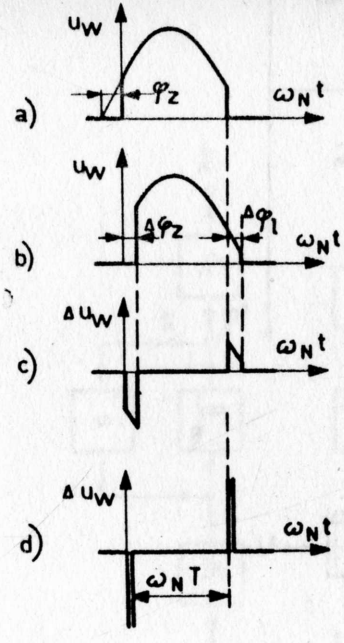


Fig 4

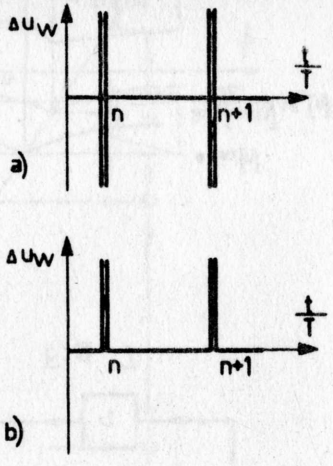


Fig 5

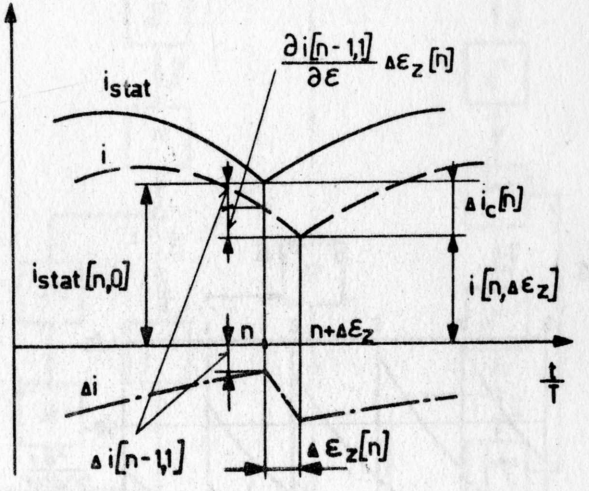


Fig 6

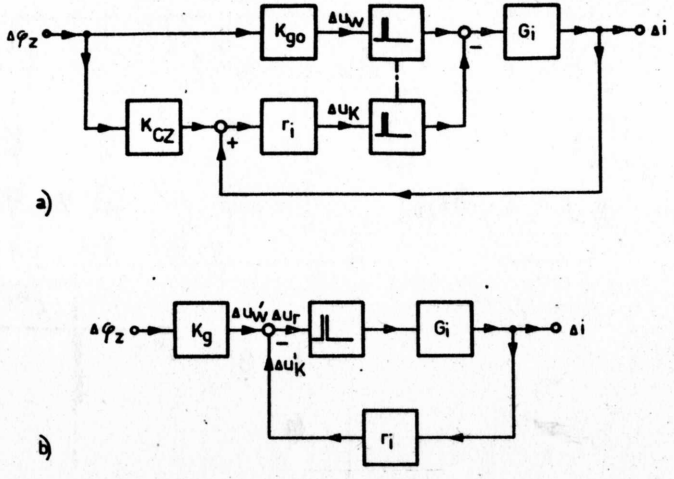


Fig 7

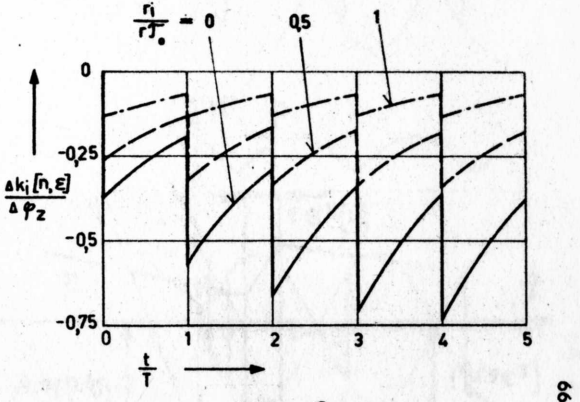


Fig 8

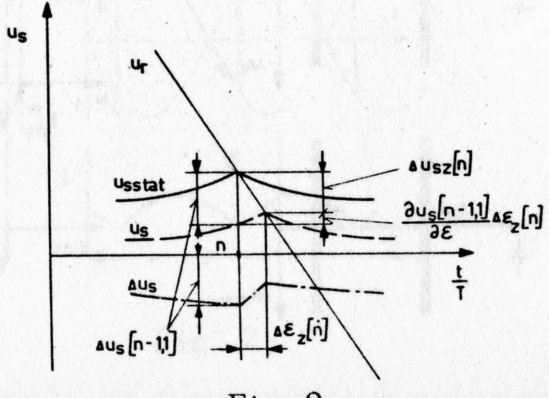


Fig 9

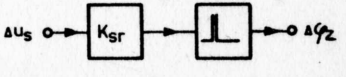
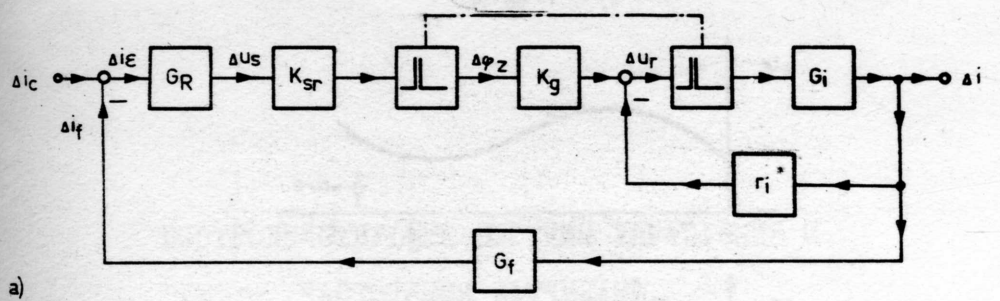
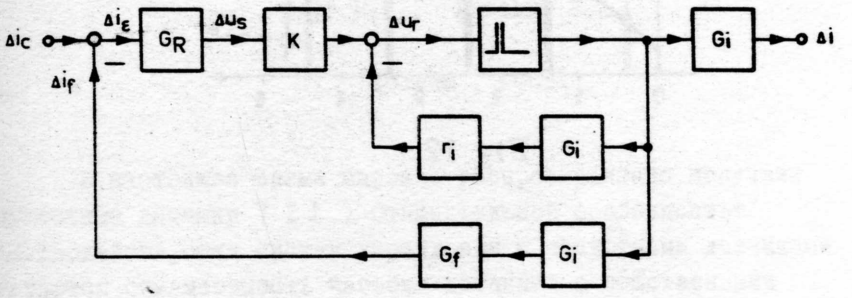


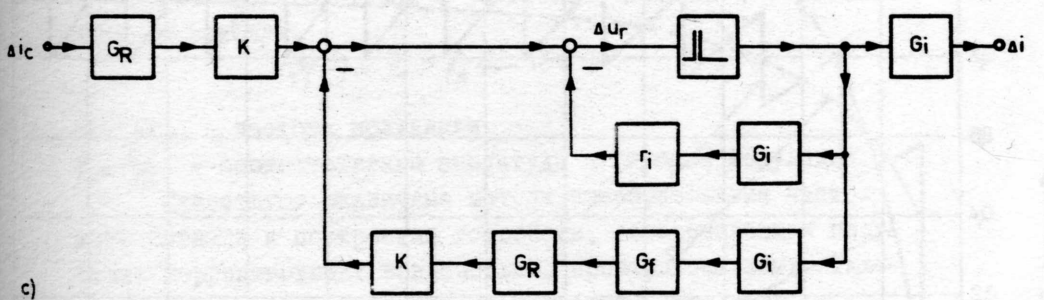
Fig 10


Δic
Δiε
GR
Δus
Ksr
Δφz
Kg
Δur
Gi
Δi
Δif
ri*
Gf

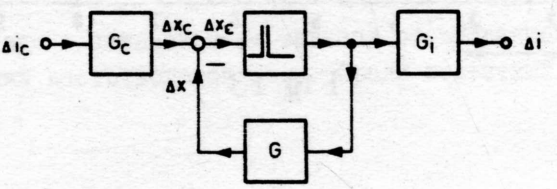
a)


Δic
Δiε
GR
Δus
K
Δur
Gi
Δi
Δif
ri
Gi
Gf
Gi

b)


Δic
GR
K
Δur
Gi
Δi
ri
Gi
K
GR
Gf
Gi

c)


Δic
Gc
Δxc
Δxc
Gi
Δi
Δx
G

d)

Fig 11

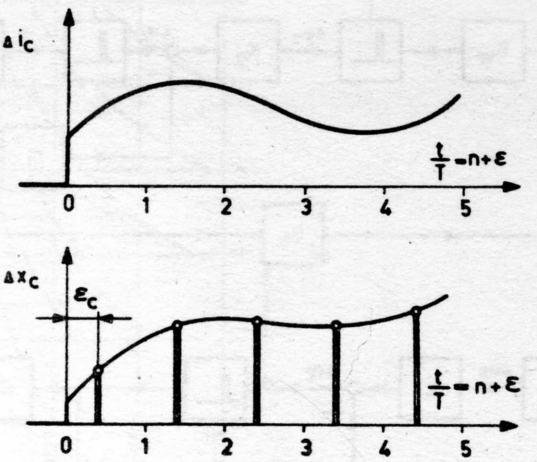

Δ i_C
t/T = n+ε
0 1 2 3 4 5
Δ x_C
ε_C
t/T = n+ε
0 1 2 3 4 5

Fig 12

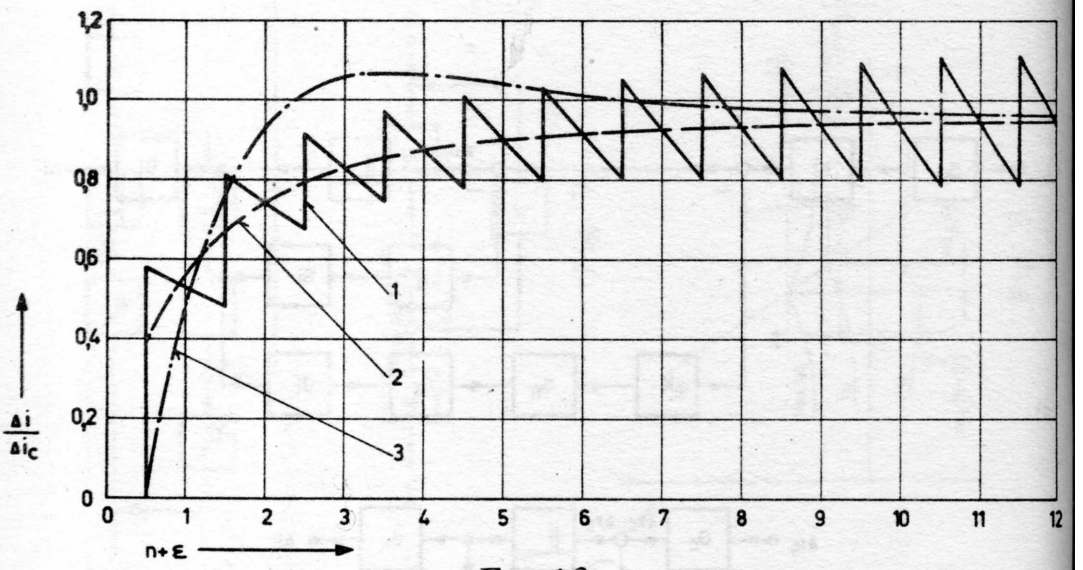

1,2
1,0
0,8
0,6
0,4
0,2
0
Δi / Δi_C
1
2
3
0 1 2 3 4 5 6 7 8 9 10 11 12
n+ε

Fig 13

БЫСТРОДЕЙСТВУЮЩИЕ СИСТЕМЫ УПРАВЛЕНИЯ С ЧАСТОТНЫМИ ДАТЧИКАМИ.

Круг Е.К., Легович Е.А.

Москва , СССР.

В настоящее время широкое распространение получили частотные датчики [I] . Отличительной особенностью быстродействующих систем управления с частотными датчиками является соизмеримость частоты датчиков с собственными частотами систем. При построении таких систем возникает ряд трудностей, связанных с воспроизведением сигналов, пропорциональных первой и более высоким производным от входного сигнала.

Примем, что сигнал с частотного датчика $f(t)$ изменяется по закону:

$$f(t) = f_0 + f_A \sin \Omega t$$

где Ω - частота модуляции

f_A и f_A/f_0 - соответственно амплитуда и глубина модуляции.

Существуют различные методы преобразования частотного сигнала и построения устройств, обеспечивающих получение корректирующих воздействий, пропорциональных отклонению частоты от её заданного значения, интегралу от отклонения, первой, второй и т.д. производным от отклонения.

Эти методы в сильной мере зависят от выбранного способа измерения частотного сигнала. Среди известных спосо-

бов измерения частотного сигнала можно выделить 2 основные группы, внутри которых могут быть различные модификации.

I. Непрерывные (аналоговые) способы измерения, при которых измерение частоты производится непосредственным детектированием частотного сигнала.

2. Дискретные (цифровые) способы, при которых измерение частоты производится или:

а) путем подсчета числа импульсов измеряемой частоты за определенный интервал времени θ . Назовем этот способ интегральным

или: б) путем подсчета числа импульсов эталонной частоты за период измеряемой частоты $T_t = \frac{1}{f(t)}$. При этом осуществляется измерение периода входной частоты.

Системы, построенные в соответствии с последним методами используют цифровую технику и при этом может быть обеспечена высокая статическая точность. Рассмотрим возможность получения корректирующих воздействий в соответствии с указанными способами измерения частоты.

Непрерывный метод преобразования частотного сигнала.

Среди существующих схем измерения частотного сигнала при сравнительно небольшой глубине модуляции ($f_1/f_0 < 0,1$) большое распространение получили схемы с частотными демодуляторами. В этих схемах частота преобразуется в аналоговую величину - ток или напряжение. Как известно, схема частотного демодулятора содержит (рис.Ia) 2 резонансных контура $Ф_1$ и $Ф_2$, настроенных на граничные частоты заданного диапазона частот, выпрямитель со сглаживающим фильтром СФ и выходной усилитель постоянного тока У с коэффициентом усиления К.

Настраиваются демодуляторы таким образом, что выходной сигнал демодулятора пропорционален отклонению частоты от её заданного значения $f_з$ (примем, что $f_з = f_0$).

Точность такой схемы при тщательной настройки может достигать 0,5 - I%.

Схема с частотным демодулятором обладает инерционностью и эта инерционность определяется в основном постоянной времени элементов сглаживающего фильтра $T_{сф}$

В таблице I во втором столбце приведены частотные характеристики различных корректирующих устройств, использующих частотный демодулятор. В этих устройствах для получения сигнала производной на выходе частотного демодулятора включают дифференцирующий элемент, состоящий из цепочки $R_1 C_1 = T_1$ и усилитель с коэффициентом K_1 ($T_{g_1} = K_1 T_1$). Соответственно для получения сигнала второй производной последовательно включают 2 таких элемента (рис.Iб). Преобразование частоты с помощью частотного демодулятора практически не может быть использовано в системах, в которых частота $f(t)$ соизмерима с частотой модуляции Ω из-за значительных фазовых погрешностей, вызванных фильтрующим элементом.

Например, если мы зададимся условием, что на выходе преобразователя амплитуда колебания частоты $f(t)$ не должна превышать I% ($2\pi f_0 T_{c\phi} = 100$), а соотношение $\frac{2\pi f_0}{\Omega}$ равно 0,0I, то фазовая погрешность ($arctg\, \Omega T_{c\phi}$) будет превышать 40%. Задаваясь величиной фазовой погрешности в 10^0 ($\Omega T_{c\phi} \leq 0{,}17$) получим, что соотношение $2\pi f_0/\Omega \geq 570$

Исследования показали [2], что при соизмеримости частот f и Ω и небольшой глубины модуляции $f_0/f_0 < 0{,}1$ целесообразно использовать умножители частот с коэффициентом умножения $N > 100$ (рис.Iб). В этом случае инерционность фильтров может быть уменьшена в N раз. Если $N \geq 100$, то постоянной времени сглаживающего фильтра можно пренебречь и динамические свойства преобразователя практически определяются запаздыванием, вносимым умножителем, равным $\frac{\Omega}{f_0}$. Соответственно частотные характеристики корректирующих устройств, использующих частотный умножитель, приведены в третьем столбце таблице I.

Такой преобразователь практически может работать в диапазоне частот $0 < \frac{\Omega}{f_0} < 0{,}2$, при этом его фазовая погрешность не превысит $5^0 - 6^0$ (например, при $f_0 = 50$ гц, $\frac{\Omega_{макс}}{2\pi} < 1{,}5$ гц).

Построение корректирующих устройств, определяющих первую и вторую производные, а также интеграл от отклонения, связано с общеизвестными трудностями, свойственными аналоговым способам преобразования. Точность аналоговых интеграторов не превышает 0,5 - I%, та же точность

свойственна аналоговым усилителям.

Практически при использовании усилителей и дифференцирующих цепочек можно в ограниченном диапазоне частот получить сигналы, пропорциональные первой и второй производным. Максимальные значения постоянных дифференцирования определяются допустимыми фазовыми искажениями и максимальным коэффициентом усиления K_1, который можно получить при $(\Omega T_1)_{макс} < 0,2$ или $(\Omega T_g)_{макс} < 0,2K_1$. При $K_1 \geq 20$ $T_{g\,макс} < \frac{4}{\Omega_{макс}}$

Таким образом, можно утверждать, что непрерывный (аналоговый) способ преобразования может быть использован при соизмеримости частот f и Ω только при применении умножителя частот с коэффициентом умножения $N > 100$

Статическая точность рассматриваемых корректирующих устройств может быть доведена до 0,5 - 1%.

Величина же фазовой погрешности, вызванная соизмеримостью частот определяется величиной $\frac{\Omega}{f_0}$. Корректирующие устройства, обеспечивающие получение сигналов пропорциональных первой и второй производных, имеют дополнительные фазовые погрешности, свойственные аналоговым методом дифференцирования.

Интегральный метод преобразования частотного сигнала.

Высокую точность и помехоустойчивость при измерении частоты можно получить при счете импульсов частоты за определенный интервал времени θ в дискретные моменты времени nT. Соответствующее выражение при таком методе измерения имеет вид

$$f_{nT} = \int_{nT}^{nT+\theta} f(t)\,dt$$

Корректирующее устройство, реализующее заданный алгоритм, содержит счетчик импульсов Сч, логический элемент ЛЭ и генератор импульсов ГИ (рис.2а). Генератор импульсов обеспечивает подачу в моменты времени nT на логический элемент напряжения в течение интервала θ, при котором на счетчик поступают импульсы частоты $f(t)$.

Частотные характеристики различных корректирующих устройств при таком методе измерения приведены в таблице 2 во втором столбце.

Мы видим, что всем системам с интегральным методом преобразования частотного сигнала свойственны значительные фазовые погрешности, вызванные временем измерения θ. Кроме того, устройства, обеспечивающие получение сигнала отклонения частоты от её заданного значения, и устройства, вычисляющие производные, имеют фазовые погрешности, определяемые квантованием по времени T [3].

Имеет также место зависимость амплитуд выходных сигналов отдельных корректирующих устройств от величины T. Из-за значительных фазовых погрешностей интегральный метод преобразования можно рекомендовать, когда частоты f и Ω не соизмеримы.

Действительно, для получения высокой точности измерения обычно принимают:

$$f_A \cdot \theta \geq 1000, \quad \Omega T \leq 0{,}1 \quad и \quad T = \theta$$

При этих условиях соотношения между f_A и Ω должны удовлетворять условию $f_A \geq 10000\,\Omega$.

При небольшой глубине модуляции ($\frac{f_A}{f_0} < 0{,}1$) для сокращения времени θ целесообразно использовать умножители частот - УЧ (рис.2а). В этом случае (как видно из таблицы 2 третий столбец) можно сократить время измерения в N раз, практически доводя время измерения до $\theta' = \frac{1}{\delta N f_A}$ где δ - погрешность измерения частоты.

В пропорциональных устройствах также можно уменьшить время цикла T, доводя его до величины θ'. Однако в системах с корректирующими устройствами, обеспечивающими получение сигналов, пропорциональных производным нельзя произвольно менять время T. Отличительной особенностью таких устройств является необходимость выбора временного интервала T в зависимости от возможного диапазона изменения частоты модуляции Ω. Проведенный анализ позволил получить конкретные соотношения между амплитудой входного сигнала $A_n = f_A \theta$, амплитудной и фазовой погрешностью выходного сигнала и величиной Ω T.

На рис.3 в логарифмическом мастабе представлены соответствующие зависимости. Сплошными линиями показаны зависимости величины ΩT от амплитуды A_n на входе корректирующего звена, воспроизводящего сигнал, пропорциональный I-ой производной, при котором амплитуда на выходе этого звена равна $A_{g_1}=20$, $A_{g_1}=50$ и $A_{g_1}=100$ квантам. Пунктирными линиями соответственно изображены зависимости ΩT от A_n при тех же значениях амплитуд на выходе корректирующего звена, воспроизводящего сигнал 2-ой производной ($A_{g_2}=20$, $A_{g_2}=50$, $A_{g_2}=100$).

Из рассмотрения рис.3 и данных таблицы 2 видно, что квантование по уровню, свойственное всем цифровым системам, ограничивает частотный диапазон ΩT снизу, а допустимая фазовая погрешность, вызванная квантованием по времени, для данного корректирующего звена ограничивает этот диапазон сверху. Кроме того, чем выше порядок корректирующего воздействия, тем уже диапазон частот и выше требования к точности измерения и точности настройки параметра T.

Например, при измерении отклонения с точностью 0,1% $f_A\theta'$ = 1000 квантам, первая производная может быть определена с точностью 5% (A_{g_1} =20 квантам) при фазовой погрешности 11^0 ($\Omega T < 0,2$) в диапазоне $\Delta(\Omega T)_1 (0,02<\Omega T<0,2)$, а вторая производная с точностью 5% при фазовой ошибки в 18^0 ($\Omega T = 0,2$) может быть получена только при $f_A\theta' \geqslant 10000$ в диапазоне $\Delta(\Omega T)_2$ ($0,14<\Omega T<0,2$).

Таким образом, можно утверждать, что при цифровом методе измерения точность получения производной первого порядка ограничена, а получение производных высоких порядков (выше второго) практически невозможна.

Однако следует отметить, что получение корректирующих воздействий пропорциональных отклонению частоты от её заданного значения и интегралу от отклонения не встречает затруднений. Интегральный цифровой метод при вычислении этих воздействий может обеспечить очень большую точность.

Преобразование частотного сигнала при измерении его периода

При заполнении периода измеряемой частоты эталонной частотой $f_э$ можно также получить высокую точность измерения. Структурная схема такого устройства представлена на рис.2б. Устройство содержит счетчик импульсов Сч, и логический элемент ЛЭ, и генератор эталанной частоты ГИ. Схема управления логическим элементом обеспечивает подачу на счетчик импульсов частоты $f_э$ в течение интервала $T_t = \frac{1}{f(t)}$. В результате в счетчике оказывается записано число $N^x = [f_э T_t]^x$.

Частотные характеристики отдельных корректирующих устройств при таком методе измерения приведены в таблице 3.

Левые столбцы таблицы относятся к случаю, когда измеряется последовательно каждый период и разности вычисляются между соседними периодами (отклонение $\Delta x_n = [f_э T_n]^x - [f_э T_{n-1}]^x$

В правых столбцах таблицы приведены данные для устройств, работающих дискретно с постоянным временем цикла Т.

Следует отметить, что системы, основанные на измерении периода, являются нелинейными. Поэтому приведенные данные практически справедливы при небольшой глубине модуляции ($f_A/f_0 < 0{,}1$) или при обеспечении линеаризации.

Мы видим, что метод измерения последовательно каждого периода является наиболее быстродействующим и может быть успешно использован для получения сигнала отклонения. Однако практически его нельзя применять при вычислении производных. Точность определения производных очень мала, так как мала сама разница между соседними периодами частоты, а амплитуды при производных соответственно пропорциональны $\frac{\Omega}{f(t)}$.

Например, если измерение производится с точностью 0,I%, что соответствует $f_э/f_0$ =I000, а входной сигнал имеет глубину модуляции $f_A/f_0 < 0{,}1$ и $\frac{\Omega}{f_A} < 0{,}1$ (например $f_0 = 5$ гц, $f_A = 5$ гц и $\frac{\Omega}{2\pi} = 0{,}5$ гц), то амплитуда отклонения A_n соответствует I00 квантам, а амплитуда первой производной не будет превосходить одного кванта ($A_{g_1} = A_n \frac{\Omega}{f_0}$ =I квант),

т.е. в такой системе даже невозможно гарантировать точность определения знака производной.

Вычисление производных целесообразно производить при введении дискретности по времени. В этом случае время цикла Т необходимо выбирать в соответствии с частотным диапазоном входного сигнала согласно графикам, приведенным на рис.3.

Практически при измерении периода при соответствующем выборе величины Т, возможно построение цифровых корректирующих устройств, вычисляющих первую производную. В этом случае, так же как и при интегральном методе, частотный диапазон входного сигнала, при котором может быть обеспечена заданная точность получения производной, ограничен снизу квантованием по уровню (величиной $\frac{f_д}{f_3}$) и сверху фазовой ошибкой, вызванной квантованием по времени.

Линеаризация при построении таких корректирующих устройств не обязательна, так как точность получения сигналов, пропорциональна производным не велика.

Следует отметить, что отсутствие линеаризации в системах с интегральным воздействием может привести к статической ошибки, если измерение T_t производится в дискретные моменты времени nT , а не каждый период частоты $f(t)$. При измерении последовательно каждого периода интегральное корректирующее устройство может рассматриваться как непрерывный интегратор, уравнение которого может быть записано в виде:

$$y = \int [f(t) - f_3]\,dt$$

Соответственно структурная схема такого устройства приведена на рис.2в. Она состоит из счетчика Сч, логического элемента ЛЭ - (элемента несовпадения) и задающего генератора ГЗ, обеспечивающего сигнал, пропорциональный заданному значению частоты f_3 .

Таким образом, на основании проведенного анализа можно сформулировать следующие рекомендации по выбору метода построения устройств, обеспечивающих получение различных корректирующих воздействий в системах при частотном входном сигнале.

I. Для получения сигнала отклонения частоты от её заданного значения можно воспользоваться всеми рассмотренными методами.

I. Непрерывный метод, основанный на применении умножителя частот и частотного демодулятора, позволяет обеспечить статическую точность определения частоты в 0,5- I%, при фазовой погрешности в I0-I2^0 в диапазоне частот $0<\frac{\Omega}{f_0}<0,1$ при глубине модуляции $f_д/f_0<0,1$.

2. Высокую статическую точность ($\delta<0,001$) можно получить как при интегральном методе измерения частоты, так и при измерении периода $T_t=\frac{1}{f(t)}$. Второй способ является бо́лее быстродействующим. Если осуществлять измерение последовательно каждого периода, то практически его можно использовать в диапазоне $0<\frac{\Omega}{f_0}<0,1$ при относительно большой величине модуляции. Однако для получения высокой статической точности необходимо предусмотреть соответствующую линеаризацию.

3. Интегральный метод может быть применен при соизмеримости частот $f(t)$ и Ω только при применении частотных умножителей с коэффициентом $N>100$ и небольшой глубине модуляции ($f_д/f_0<0,1$). Практически частотный диапазон лежит в пределах $0<\frac{\Omega}{f_0}<0,1N\delta$.

П. Для получения сигнала, пропорционального первой производной при соизмеримости частот $f(t)$ и Ω целесообразно использовать как непрерывные дифференциаторы в сочетании с умножителем частоты и частотным демодулятором, так и быстродействующие цифровые дискретные дифференциаторы. Оба эти метода примерно равноценны с точки зрения статической точности и динамической ошибки.

I. При аналоговом методе дифференцирования и при использовании умножителя частоты с демодулятором диапазон частоты T,Ω ограничен величинами $0<\Omega T,<0,1$ или $0<\frac{\Omega}{f_0}<0,1$. Статическая точность такого дифференциатора порядка I -2 %. Динамическая погрешность на превышает I8^0.

2. Быстродействующий цифровой дифференциатор целесообразно строить при $0<\frac{\Omega}{f_0}<0,1$ как устройство, вычисляющее разность между двумя периодами частоты $f(t)$, измеренными в моменты времени nT. Необходимо предусмотреть настройку времени цикла T в зависимости от диапазона изменения,

Цифровой дифференциатор может работать в диапазоне $0{,}02 < \Omega T < 0{,}2$, где нижняя граница определяется квантованием по уровню, а верхняя граница допустимой фазовой погрешностью.

Ш. Получение сигнала, пропорционального второй производной, связано с большими трудностями при использовании как цифровых, так и аналоговых методов. Точность обеспечения такого сигнала не высока. Фазовые погрешности значительны. Выбор того или другого метода зависит от устройств, используемых при вычислении отклонения и первой производной.

Вне зависимости от соизмеримости частот $f(t)$ и Ω построения цифровых корректирующих устройств, обеспечивающих вычисление производных выше второй, не имеет смысла.

IУ. Для получения сигнала пропорционального интегралу от ошибки целесообразно применять цифровой интегральный метод измерения частоты. При этом обеспечивается высокая точность измерения и регулирования при использовании как непрерывного, так и дискретного интегратора[4].

I. Цифровой непрерывный интегратор позволяет работать в диапазонах частот $0 < \frac{\Omega\theta}{2} < 0{,}2$ практически при $f_д/f_0 < 0{,}5$.

2. Цифровой дискретный интегратор в диапазоне $0 < \frac{\Omega\theta}{2} < 0{,}2$ при $\theta < T$

Интегральный метод можно также рекомендовать для построения рассмотренных корректирующих устройств в случае, если частоты $f(t)$ и Ω не соизмеримы.

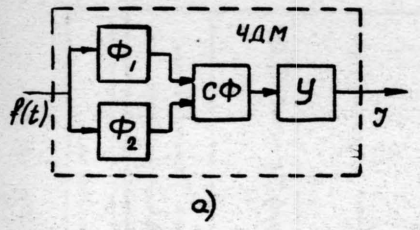
Литература.

I. Агейкин Д.М. и др. "Датчики контроля и регулирования" "Машиностроение" I965 г.

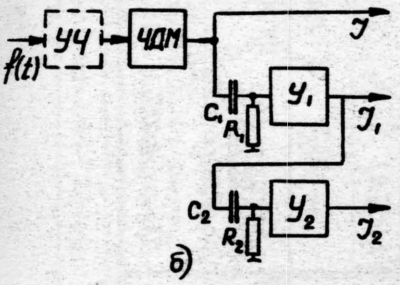
2. Фатеева Е.А, "К вопросу о построении преобразователей низкой частоты" А и Т №2 I967г

3. Круг Е.К. "Некоторые вопросы применения цифровых методов коррекции" А и Т №8 I967г.

4. Круг Е.К., Александриди Т.М., Дилигенский С.Н. "Цифровые регуляторы" "Энергия" I966г.

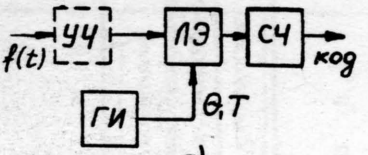


а)

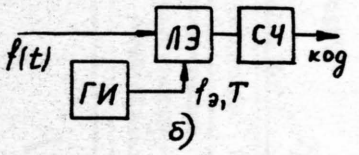


б)

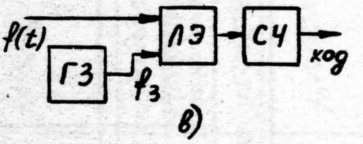
Рис.1



а)



б)



в)

Рис.2

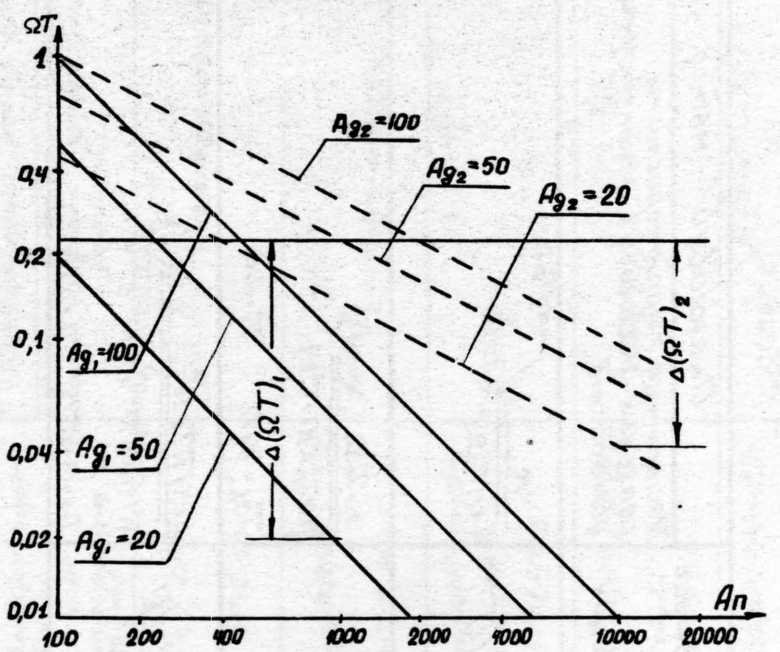


Рис.3

Таблица №1

Идеализированные уравнения аналоговых корректирующих устройств	Аналоговый метод преобразования частотного сигнала	
	Частотные характеристики реальных аналоговых корректирующих устройств.	Частотные характеристики реальных аналоговых корректирующих устройств с предварительным умножением.
$y(t) = K_n x(t)$	$\frac{K}{\sqrt{1+(T_{\phi o}\Omega)^2}} e^{-j\,\mathrm{arctg}\,T_{\phi o}\Omega}$	$K e^{-j\left(\frac{\Omega}{f_o}\right)}$
$y(t) = T_{g1}\frac{dx(t)}{dt}$	$\frac{K_1 T_1 \Omega}{\sqrt{(1+T_{\phi o}^2\Omega^2)(1+T_1^2\Omega^2)}} e^{j\left(\frac{\pi}{2}-\mathrm{arctg}\,T_{\phi o}\Omega-\mathrm{arctg}\,T_1\Omega\right)}$ $T_{g1}=K_1T_1 \qquad T_1=R_1C_1$	$\frac{K_1 T_1 \Omega}{\sqrt{(1+T_1\Omega)}} e^{j\left(\frac{\pi}{2}-\frac{\Omega}{f_o}-\mathrm{arctg}\,T_1\Omega\right)}$ $T_{g1}=K_1T_1 \qquad T_1=R_1C_1$
$y(t) = T_{g2}^2\frac{d^2x(t)}{dt^2}$	$\frac{K_1K_2T_1T_2\Omega^2}{\sqrt{(1+T_{\phi o}^2\Omega^2)(1+T_1^2\Omega^2)(1+T_2^2\Omega^2)}} e^{j(\pi-\mathrm{arctg}\,T_{\phi o}\Omega-\mathrm{arctg}\,T_1\Omega-\mathrm{arctg}\,T_2\Omega)}$ $T_{g2}=K_1K_2T_1T_2 \qquad T_2=R_2C_2$	$\frac{K_1K_2T_1T_2\Omega^2}{\sqrt{(1+T_1^2\Omega^2)(1+T_2^2\Omega^2)}} e^{j\left(\pi-\frac{\Omega}{f_o}-\mathrm{arctg}\,T_1\Omega-\mathrm{arctg}\,T_2\Omega\right)}$ $T_{g2}=K_1K_2T_1T_2 \qquad T_2=R_2C_2$

Таблица N°2

Идеализированные уравнения дискретных корректирующих устройств	Интегральный метод преобразования частотного сигнала	
	Частотные характеристики по 1й гармонике реальных дискретных корректирующих устройств.	Частотные характеристики по 1й гармонике реальных дискретных корректирующих устройств с предварительным умножением частоты.
$y(t)=K_n\sum_{1}^{n}x[nT]$ при $nT<t<(n+1)T$	$\frac{f_A\cdot\Theta}{\Omega T}\,e^{j\left(-\frac{\pi}{2}-\frac{\Omega}{2}\right)}$	$\frac{f_A N\Theta'}{T\Omega}e^{-j\frac{\pi}{2}}$ $\Theta'=\frac{\Theta}{N}$
$y(t)=K_n x[nT]$ при $nT<t<(n+1)T$	$f_A\cdot\Theta\,e^{j\left(-\frac{T\Omega}{2}-\frac{\Theta\Omega}{2}\right)}$	$f_A\cdot N\Theta'\,e^{-j\frac{T\cdot\Omega}{2}}$
$y(t)=K_1(x[nT]-x[(n-1)T]=$ $=K_1\Delta_1 x[nT]$ при $nT<t<(n+1)T$	$f_A\Theta\cdot T\Omega\;e^{j\left(\frac{\pi}{2}-\Omega T-\frac{\Theta\Omega}{2}\right)}$	$f_A N\Theta\cdot\Omega T\,e^{j\left(\frac{\pi}{2}-\Omega T\right)}$
$y(t)=K_2(\Delta x[nT]-\Delta x[(n-1)T])=$ $=K_2\Delta_2 x[nT]$ при $nT<t<(n+1)T$	$f_A\cdot\Theta(T\Omega)^2\,e^{j\left(\pi-\frac{3}{2}T\Omega-\frac{\Theta\Omega}{2}\right)}$	$f_A N\Theta'(\Omega T)^2\,e^{j\left(\pi-\frac{3}{2}\Omega T\right)}$
$y(t)=K_3(\Delta_3 x[nT]-\Delta_2 x[(n-1)T])=$ $=K_3\Delta_3 x[nT]$ при $nT<t<(n+1)T$	$f_A\Theta(T\Omega)^3\,e^{j\left(\frac{3}{2}\pi-2T\Omega-\frac{\Theta\Omega}{2}\right)}$	$f_A N\Theta'(\Omega T)^3\,e^{j\left(\frac{3}{2}\pi-2\Omega T\right)}$

Таблица №3

Вид корректирующего сигнала.	Преобразование частотного сигнала при измерении его периода.	
	Частотные характеристики по 1ой гармонике реальных дискретных корректирующих устройств при измерении каждого периода $f(t)$.	Частотные характеристики по 1ой гармонике реальных дискретных корректирующих устройств при измерении периода $f(t)$ через время цикла T.
Интеграл	$\frac{f_э f_д}{f_0^2\left(\frac{\Omega}{f_0}\right)} e^{j\left(-\frac{\pi}{2}-\frac{\Omega}{2f_0}\right)}$	$\frac{f_э f_д}{f_0^2 \Omega T} e^{-j\frac{\pi}{2}}$ при $\frac{f_д}{f_0} \angle 0{,}1$
Пропорциональный отклонению	$\frac{f_э f_д}{f_0^2} e^{-j\frac{\Omega}{2f_0}}$	$\frac{f_э f_д}{f_0^2} e^{-j\frac{T\Omega}{2}}$
1ая производная	$\frac{f_э f_д}{f_0^2}\frac{\Omega}{f_0} e^{j\left(\frac{\pi}{2}-\frac{3\Omega}{2f_0}\right)}$	$\frac{f_э f_д}{f_0^2} \Omega T e^{j\left(\frac{\pi}{2}-\Omega T\right)}$
2ая производная	$\frac{f_э f_д}{f_0^2}\left(\frac{\Omega}{f_0}\right)^2 e^{j\left(\pi-\frac{2\Omega}{f_0}\right)}$	$\frac{f_э f_д}{f_0^2}(\Omega T)^2 e^{j\left(\pi-\frac{3}{2}\Omega T\right)}$
3ья производная	$\frac{f_э f_д}{f_0^2}\left(\frac{\Omega}{f_0}\right)^3 e^{j\left(\frac{3}{2}\pi-\frac{5\Omega}{2f_0}\right)}$	$\frac{f_э f_д}{f_0^2}(\Omega T)^3 e^{j\left(\frac{3}{2}\pi-2\Omega T\right)}$

OPTIMUM CALIBRATION OF INERTIAL COMPONENTS

George T. Schmidt
Instrumentation Laboratory, Massachusetts Institute of Technology, Cambridge, Mass.

Introduction

Two recent developments are leading to changes in the way inertial instruments are calibrated in component, system, and vehicle level tests. One of these developments is the new statistical filtering, prediction, and smoothing techniques which provide the theory to account for all random noises, measurement errors, and disturbances affecting the component under test and its environment. The other complementary development is the availability of more powerful computers, both the large digital machine and the vehicle on-board computer, which provide the computational capability to take advantage of modern time-variable filter theory. There are, however, several practical problems encountered when one attempts to implement these techniques. A major one is in obtaining an adequate statistical model of the physical systems involved in the calibration process. Another problem is the immense amount of computation required. These two problems are interrelated in that the more accurately the system is statistically modeled, the more difficult the computational problem becomes.

The plan for this paper is to first show the form of any calibration system that uses statistical filtering together with the basic filter equations. Two applications with solutions will then be presented. The first is the use of a large digital computer to process gyro test table data and provide knowledge of gyro drift together with generating a model of gyro behavior. The filter is formulated and actual gyro test data is processed; the results are compared with a conventional Fourier series data reduction. This problem poses very little computational difficulty because of the availability of a large digital computer.

The second application is to the alignment and calibration of an inertial platform in a launch vehicle that is subjected to wind induced sway while on the launch pad. This application requires an engineering solution to the real time computation problem with a limited on-board computer. We also impose the requirement that the calibration procedure be independent of any external references. Only the outputs of the components on the platform are to be used for calibration. The problem solution is discussed in detail and is, in fact, the method used in the Apollo Guidance Navigation and Control System.

Form of a Calibration System Using Statistical Filtering

The basic configuration of an inertial component calibration system is shown in functional form in Fig. 1. The question is, "What should the computer program be that converts the measurements to the desired outputs?" In other words, given a set of noisy physical measurements and some knowledge of the statistical characteristics of the components involved (e. g., accelerometers on a gyro-stabilized platform), find the best estimate of the desired output quantities.

The maximum amount of information provided by the measurements in a system with unpredictable errors is the probability density function $p(\underline{x})$. The vector $\underline{x}$ represents the state of the entire system involved including the sensors, the vehicle(or mounting for the sensors), and the environment. With the knowledge of all past measurements and the a priori assumptions of the statistical characteristics of the entire system, the most that can be known about the state of a system is the probability that it is in each incremental volume of possible states. In most cases the entire probability density function is not desired. What is usually wanted is a set of values that is, in some respect, the "best" estimate of the state of the system. This estimate is symbolized by $\hat{\underline{x}}$, and it is usually chosen so as to minimize the average of some function of the error in the estimate $\underline{e} = \hat{\underline{x}} - \underline{x}$. The most obvious estimate and one that minimizes the mean-squared error is the expected or mean value which is given by

$$\hat{\underline{x}} = E(\underline{x} \,|\, \underline{m}) = \int_{-\infty}^{\infty} \underline{x}\, p(\underline{x} \,|\, \underline{m}_1, \ldots, \underline{m}_n)\, d\underline{x} \qquad (1)$$

where $p(\underline{x} \,|\, \underline{m}_1, \ldots, \underline{m}_n)$ is the probability density function conditioned by all past measurements.

The solution of Eq. (1) would require the computation of the entire probability density function of all the state variables as a function of time. The computation would include the changes in the density function due to the dynamics of the system and the changes due to the information gained by new measurements. The techniques for accomplishing this computation are not, to the author's knowledge, well developed and, furthermore, would be impractical for any computer in the foreseeable future. However, if we can assume that the errors in the estimate can be represented by a linear system excited by uncorrelated noise, then the whole process represented by Eq. (1) can be stated concisely by equations developed by Kalman.[1, 2]

Mathematical Development

It is assumed that the state of the entire system, including sensors, mounting, and environment, can be described by differential equations of the form

$$\dot{\underline{x}}(t) = \underline{f}\left[\underline{x}(t),\ \underline{c}(t),\ \underline{n}(t),\ t\right] \tag{2}$$

where $\underline{c}$ are known control inputs and $\underline{n}$ are white noises. For the assumptions to be made here, it can be shown that the control does not affect the form of the optimum filter.[3,4] Thus the control variables $\underline{c}$ will not be shown explicitly in the following discussion. It is assumed that measurements are made at discrete times according to the relation

$$\underline{m}(t_m) = \underline{h}\left[\underline{x}(t_m),\ \underline{u}(t_m)\right] \tag{3}$$

where $\underline{u}(t_m)$ are errors in the measurements that are uncorrelated between measurements. (Time correlated errors in the measurements would have to be included in the state vector $\underline{x}$.) Assuming that the optimal estimates are close enough to the true values so that higher order terms may be neglected, the optimum measurement process is given by Kalman's optimum linear filter. The derivations of these equations are given in Refs. 1, 2 and 4. The basic equations are

$$\left.\begin{aligned} \hat{\underline{x}} &= \hat{\underline{x}}' + E'H^T (HE'H^T + U)^{-1} (\underline{m} - \underline{h}(\hat{\underline{x}}', t)) \\ E &= E' - E'H^T (HE'H^T + U)^{-1} HE' \end{aligned}\right\} \text{at a measurement time}$$

$$\left.\begin{aligned} \dot{\hat{\underline{x}}} &= \underline{f}(\hat{\underline{x}},\ t) \\ \dot{E} &= FE + EF^T + N \end{aligned}\right\} \text{between measurements} \tag{4}$$

where the prime indicates conditions that exist just before the measurement. The covariance matrix of errors E is defined by

$$E = \langle \delta \underline{x}\ \delta \underline{x}^T \rangle$$

where "< > " represents the expected or mean value, and

$$F = \frac{\partial \underline{f}(\hat{\underline{x}}, t)}{\partial \hat{\underline{x}}} \qquad H = \frac{\partial \underline{h}(\hat{\underline{x}}, t)}{\partial \hat{\underline{x}}}$$

$$U = \frac{\partial \underline{h}}{\partial \underline{u}} R \frac{\partial \underline{h}^T}{\partial \underline{u}} \qquad N = \frac{\partial \underline{f}}{\partial \underline{n}} Q \frac{\partial \underline{f}^T}{\partial \underline{n}} .$$

The matrices R and Q are defined by

$$\langle \underline{u}(t)\, \underline{u}(\tau)^T \rangle = R\,\delta(t-\tau)$$

$$\langle \underline{n}(t)\, \underline{n}(\tau)^T \rangle = Q\,\delta(t-\tau).$$

The optimum filter for a linearized system in which the higher order terms can be neglected is shown in diagram form in Fig. 2.

Gyro Unit Testing Using Statistical Filtering

Test Configuration

The test we will consider is the so called "IA vertical" test in which the gyro with its input axis vertical is mounted on a rotary table. The gyro output axis is west and the spin reference axis south. The gyro output signal is used to drive the rotary table by means of amplifiers and motors. The table rotates about the vertical according to the differential equation:

$$da/dt = d + Wh\ b \cos(a) + Wh\ c \sin(a) + Wv \qquad (5)$$

where: a = table angle measured clockwise from west, d = gyro drift, Wv = vertical component of earth rate, Wh = horizontal component of earth rate, b = misalignment about the output axis, c = misalignment about the spin axis, and the signs of b and c are by the user's convention. Other types of servo tests are possible and are discussed in Ref. 7.

In a conventional servo table test, the time it takes for the table to move each one degree increment is recorded for two revolutions of the table; this requires about 72 hours of testing. The usual data reduction technique is to fit a Fourier series to the data and identify the constant in the series as the gyro drift and the coefficients of the cosine and sine terms as the misalignment angles. Typically the autocorrelation function is also calculated for the data and one attempts to find an autocorrelation function for some process that fits the data; eg., a ramp or perhaps a

random walk. The curve fits are usually done by a least squares method. The purpose here is to give an alternate approach using recursive statistical filtering, compare the results from both methods and the advantages in each, and suggest some areas for further research. A functional view of the optimal calibration system is shown in Fig. 3.

The Filter

The time it takes the table to go each one degree is recorded on tape. The tape is fed into a digital computer which has programmed into it the following filter.

Between each measurement the state vector is integrated according to:

$$\frac{d\hat{\underline{x}}}{dt} = \frac{d}{dt}\begin{bmatrix} a \\ d \\ b \\ c \\ e \end{bmatrix} = \begin{bmatrix} 0 & 1 & Wh\cos(a) & Wh\sin(a) & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} a \\ d \\ b \\ c \\ e \end{bmatrix} + \begin{bmatrix} Wv \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad (6)$$

where e has been introduced to account for any bias in the measurement of table angle. The initial condition for the state vector integration is the estimated value after the last measurement incorporation.

The extrapolation of the covariance matrix between measurements is a bit harder to derive. Linearizing the differential equation for a we have:

$$\delta\dot{\hat{a}} = \delta\hat{d} + Wh\cos(\hat{a})\delta\hat{b} + Wh\sin(\hat{a})\delta\hat{c} - Wh\,\hat{b}\sin(\hat{a})\delta\hat{a} + Wh\,\hat{c}\cos(\hat{a})\delta\hat{a} \quad (7)$$

The measurements of table angle - denoted by m - are assumed to be highly accurate; typically, they are true to within a few arc-seconds. Since the misalignment angles are small, one may neglect the last two terms in Eq. (7) and the F matrix in the covariance matrix extrapolation

$$dE/dt = F\,E + E\,F^{T} \quad (8)$$

is given by the matrix in Eq. (6). The initial condition for the covariance matrix is the value after the last measurement incorporation. Equations (6) and (8) are integrated on the digital computer until the time of the next measurement (m). At that time the equations are updated according to

$$\underline{k} = E'\underline{h}\,(\underline{h}^T E'\underline{h} + ra)^{-1}$$
$$\hat{\underline{x}} = \hat{\underline{x}}' + \underline{k}\,(m - \underline{h}^T \hat{\underline{x}}')$$
$$E = (I - \underline{k}\,\underline{h}^T)\,E'$$

where: ra is the variance of the measurement noise and $\underline{h}^T = (1, 0, 0, 0, 1)$. The integration process (Eqs. 6 and 8) is then reinitiated.

Test Results

An "IA vertical" test was run on a gas-bearing gyro and the data recorded as described. A digital computer was programmed to filter the information. One full revolution of data in one-degree increments was processed. The initial condition for the state vector was zero and all initial cross-correlation terms were set zero. The initial diagonal of the covariance matrix was: 9×10^{-8} rad^2, 100 meru^2, 9×10^{-8} rad^2, 1×10^{-6} rad^2, and 1×10^{-10} rad^2. (A meru is approximately 0.015 degrees/hour.) The rms measurement noise was assumed to 1×10^{-5} rad.

The results of this test are plotted in Figures 4, 5 and 6. In Fig. 4 we notice, first, that the estimated drift is in two samples very close to its steady state value of 12.158 meru. The estimated rms error $\sqrt{E_{22}}$ is also plotted. In two samples it is down from its initial value of 10 meru to 0.06 meru; in 360 samples it is down to 0.003 meru.

In Fig. 5 the b estimate is plotted together with its estimated rms uncertainty. Apparently it takes about half a revolution before we have enough data to filter this quantity. The final $\sqrt{E_{33}}$ is 0.001 mrad as opposed to its initial value of 0.3 mrad. Similar comments hold for Fig. 6. The final value of $\sqrt{E_{44}}$ is 0.0007 mrad as opposed to its initial value of 1 mrad. The final angle estimates in each case are 0.489 and 0.743 mrad, respectively. Finally, the estimate of e showed an insignificant amount of bias.

Comparison with Conventional Method

The same data presented in the example was fitted with a Fourier series. The drift and misalignment angles (12.17 meru, 0.544 mrad, 0.711 mrad) compared almost perfectly with those as determined by the filter (12.158 meru, 0.489 mrad, 0.743 mrad). One might ask the question, "Why bother with this sophisticated method if a Fourier series works?"

One possible advantage is that the data can be processed as it is received. One might imagine a number of test tables tied into a central computer which would calibrate a whole number of gyros at once and provide

a best estimate of each gyro's current behavior. Another advantage is that the filter generates an rms uncertainty estimate in the estimated variables. If one has confidence in the statistics that must be assumed a priori, then in an ensemble sense, we can tell how good our test is. Furthermore, another advantage is in the ease with which additional state variables can be included in this formulation. We might like to try to add in a model for drift other than a bias; perhaps a ramp, an exponentially correlated process, or a random-walk. These models are easily added to the filter.

One possible disadvantage may be found if we test a gyro that has sudden jumps to different steady-state levels. In all likelihood, the filter will probably take a long time to catch up to the actual gyro or it may never get there. This is the usual argument for a Fourier series expansion of any discontinuous function. In any case, the purpose here is to show the form of the system that could be used to reduce gyro test table data, not that any great improvement is guaranteed. What is needed is more research into this type of formulation.

Methods for Reducing Computation

Throughout the previous example it has been tacitly assumed that a large digital computer is available to do the computation. For this simple example five state variables were required. Since the number of state variables rapidly increases with the complexity of the problem, engineering solutions will usually be required particularly if the digital computer has a limited capability. This is the case in an on-board vehicle computer. Generally there are two possible ways to reduce the computational burden.

The first method is to partition the total filter into smaller and simpler filters by neglecting the cross-correlations between dynamically unrelated variables. One might also neglect to implement minor gains for particular state variables; i.e., each state variable is updated by all measurements even if the particular measurement has an insignificant effect on that particular state variable.

The second method of simplifying the optimum filter is through the use of precomputed gains. If the system is linear, the gains at each measurement are only functions of time and the a priori assumptions of the statistics of the noises and the initial state. By specifying the measurement schedule or rate, the gains may be precomputed and stored in the vehicle's computer. For a small number of measurements this is a practical solu-

tion to the problem of implementing an optimum filter. For a large number of measurements, the precomputed gains are usually smoothly varying with respect to time and may be approximated by suitable functions that give an almost identical filter response as the true gains. The weighting function box in Fig. 2 now contains simple functions of time.

These simplification techniques will be applied to the pre-launch calibration and alignment of an inertial platform in a spacecraft on top of a swaying launch vehicle and they are, in fact, the techniques used in the Apollo Guidance Navigation and Control System. Practical hardware and software problems that were involved will also be discussed in detail.

Pre-Launch Calibration and Alignment

The inertial system to be calibrated and aligned includes gyroscopes and accelerometers. The known gravity acceleration is used to calibrate the accelerometers; the known vector rotation of gravity (earth rate) is used to calibrate the gyros. The exact quantities to be measured will not be considered at this point in developing the general optimum method of platform alignment to a local vertical coordinate system and measurement of the south and vertical gyro drifts. In this procedure, the platform is approximately aligned to the local vertical coordinates, then it goes inertial. The two horizontal accelerometer outputs (south and east) are used by the optimum filter to generate estimates of the relevant quantities by comparing the measurement of the rotation of the gravity vector with the known rotation rate.

The vertical gyro drift is the most difficult quantity to measure since it causes only a third-order effect on the measured acceleration. Gyro failures can be closely associated with changes in drift due to acceleration of gravity along the input axis so the pre-launch calibration of a gyro in a vertical position is highly desirable.

Since the estimates of the alignment and drift variables will depend on the measurement by the accelerometers of the rotation of the gravity vector in the inertial coordinates instrumented by the gyros, the major disturbances are the accelerometer quantization and the wind-induced sway of the launch vehicle. The model of the system for the optimum filter must include variables due to this sway. The complete filter must be simulated on a digital computer; it will be linear, so that by specifying the measurement schedule the optimum gains may be precomputed. The gains will be approximated by functions that will be stored in the flight computer. The

method for using this simplified filter in other platform positions, an illustration of a system test program, and practical hardware problems will be presented.

Models

The launch vehicle bending dynamics in the north - south and east-west directions are approximated by identical second-order systems. The wind causing the vehicle sway is assumed to be exponentially correlated with a correlation time of $1/\lambda$ sec. The correlation function of the white noise required to produce a mean-squared value of missile sway can be found to be[6]:

$$nw = \langle n(t)\, n(\tau) \rangle = \frac{\langle p^2 \rangle 4\lambda\zeta\omega_n^3 (\omega_n^2 + 2\lambda\zeta\omega_n + \lambda^2)}{\lambda + 2\zeta\omega_n}$$

where $\langle p^2 \rangle$ is the expected mean-squared missile sway, n(t) is the white noise generating the exponentially correlated wind, and ω_n and ζ are the natural frequency and damping ratio of the second-order approximation to the bending dynamics. The state vector for the sway variables in the south direction is

$$\begin{bmatrix} \dot{p}_s \\ \dot{v}_s \\ \dot{a}_s \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -\lambda\omega_n^2 & -\omega_n^2 - 2\zeta\lambda\omega_n & -2\zeta\omega_n - \lambda \end{bmatrix} \begin{bmatrix} p_s \\ v_s \\ a_s \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ n(t) \end{bmatrix} \qquad (9)$$

where p_s, v_s, a_s are the horizontal displacement, velocity, and acceleration in the north-south direction; the model for the sway variables in the east direction (p_e, v_e, a_e) has the same form. For the computer simulations in this paper, the preceding variables have the values: $\langle p^2 \rangle = 100\ cm^2$ (east and south), $\lambda = 0.1\ sec^{-1}$, $\omega_n = 2.09$ rad/sec, and $\zeta = 0.1$.

The orientation of the platform with respect to a local vertical coordinate system (vertical, south and east) is described by three angles (α, β, γ). If the platform axes (x, y, z) were rotated by $-\alpha$, $-\beta$, $-\gamma$, the axes would coincide with the reference coordinates. The state vector

equation for this substate is given by

$$\begin{bmatrix} \dot{\alpha} \\ \dot{\beta} \\ \dot{\gamma} \\ \dot{d}_x \\ \dot{d}_y \end{bmatrix} = \begin{bmatrix} 0 & 0 & \Omega_h & 1 & 0 \\ 0 & 0 & \Omega_v & 0 & 1 \\ -\Omega_h & -\Omega_v & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ \gamma \\ d_x \\ d_y \end{bmatrix} + \begin{bmatrix} t_x - \Omega_v \\ t_y + \Omega_h \\ t_z + d_z \\ 0 \\ 0 \end{bmatrix} \tag{10}$$

which has been assumed to be, for angle magnitudes of interest, a valid representation of the general nonlinear platform dynamics. Ω_h and Ω_v are the horizontal and vertical components of earth rate at the test site; d_x, d_y, d_z are the constant drifts for the vertical, south, and east gyros; and t_x, t_y, t_z are the torquing rates (if any) applied to the gyros. It is assumed that the torquing rates, the components of earth rate, and the east gyro drift are known perfectly, so that the vector on the right represents known forcing $\underline{c}(t)$ and is independent of the state of the system.

Using the small angle approximations, the south and east accelerometer pulse rate outputs due to platform orientation in the gravity field may be written as

$$\begin{bmatrix} \dot{po}_s \\ \dot{po}_e \end{bmatrix} = g \begin{bmatrix} -\gamma \\ \beta \end{bmatrix} \tag{11}$$

where po_s and po_e represent the total pulse counts at some instant of time; and g is the local gravity (cm/sec^2). A 1-cm/sec - pulse accelerometer quantization has been assumed.

Accelerometer pulse rates cannot be instantaneously measured but the total pulses can be counted which make up the output due to sway velocity and orientation in the gravity field. Inherent in these measurements, then, are quantization errors. The measurements are

$$\underline{m} = \begin{bmatrix} m_s \\ m_e \end{bmatrix} = \begin{bmatrix} v_s \\ v_e \end{bmatrix} + \begin{bmatrix} po_s \\ po_e \end{bmatrix} + \begin{bmatrix} n_m \\ n_m \end{bmatrix} \tag{12}$$

The term "n_m" represents the quantization error at every sampling of the accelerometer pulse count registers. Although the measurement noise is not normally distributed with zero mean(a requirement for the optimum filter), and is, in fact, uniformly distributed, the quantization operation is

assumed a normally distributed error. This viewpoint does not hinder the optimum filter's estimates of the alignment and gyro drift variables.[4] It does degrade the accuracy of the sway variables estimates, which is of little concern. Also

$$\langle n_m(t)\, n_m(t+\tau)\rangle = ra\ \delta(\tau) \tag{13}$$

The measurements are assumed to be made every second.

The state vector is 13-dimensional; the state vector differential equation is

$$d\underline{x}/dt = F\underline{x} + \underline{c} + \underline{n}(t) \tag{14}$$

(See Fig. 7 for $\underline{x}$, F, $\underline{c}$, and $\underline{n}$). With the derivation of the model for the system, the complete optimum linear filter is defined. The accelerometer pulse count registers will be sampled at constant rates. The estimated state vector is extrapolated between measurements according to

$$\hat{\underline{x}}' = \Phi\hat{\underline{x}} + \underline{c} \tag{15}$$

and the covariance matrix according to

$$E' = \Phi E\,\Phi^T + S \tag{16}$$

where Φ and S are precomputed constant matrices for the time step between measurements. They satisfy the following differential equations:

$$dS/dt = FS + SF^T + N \qquad S(0) = 0 \tag{17}$$

$$d\Phi/dt = F\Phi \qquad \Phi(0) = I \tag{18}$$

which may be integrated on a digital computer for a time step between measurements of 1 sec.[5] At the time of a measurement, $\hat{\underline{x}}$ and E are changed according to Eq. (4). (See Fig. 7 for H, U, and N.)

Filter Design

Computer Simulations

A complete nonlinear simulation of the inertial platform in a swaying launch vehicle was made on a digital computer in order to simulate real accelerometer outputs. The initial misalignments were 1 degree on all axes; drifts were 10 meru for the vertical and south gyros and zero for the east gyro. The rms sway in both horizontal directions was 10 cm. The initial conditions for the covariance matrix were 1 deg^2 for the alignment angles and 100 $meru^2$ for the gyro drifts. All initial cross correlation terms were assumed zero. The initial estimate of the state was a zero vector.

The response of the filter was excellent for these conditions. A number of runs were first made in which the matrix U was varied so as to cause good agreement between the rms error as determined by the filter and the actual error. The errors between the estimates and the actual values of azimuth angle, vertical gyro drift and south gyro drift are shown in Figs. 8a, b and c. The errors reach small values for the three cases in 15, 40 and 10 min respectively. The errors in the estimates of the two leveling angles (β and γ) are negligible after the first few measurements. The estimates of the sway variables are not particularly good but this is not important. These simulations were run under assumed perfect knowledge of east gyro drift because of the classical result that east gyro drift cannot be identified from azimuth error. The fact that east gyro drift must be known presents no problem; as seen from Fig.8c, the south gyro can be calibrated in about 10 minutes and the error shows little sensitivity to d_z. This gyro can then be placed east and a complete calibration and alignment made. (The question of other platform positions is discussed later.)

Design of the Simplified System

The gains for the optimum filter may be precomputed for all trials since the measurement times will be the same and the a priori assumption for the statistics of the initial state vector and noises will not change. For the problem at hand, the implementation of the gains into the system involved, first, the design of a simplified optimum filter. The gains for each state variable depend on both accelerometer measurements and, in general, one gain is much smaller than the other and can be neglected. In this problem all cross-coupling measurement gains are neglected;

e. g., vertical drift estimation depends primarily on the south accelerometer so the east accelerometer measurement gain for vertical drift is not implemented. Typically the predominant gains vary as in Fig. 9. These gains can be approximated by exponentials and straight-line segments where, at distinct intervals, the time constants and slopes are changed to continually fit the approximate gains to the true gains. The gains for the six sway variables quickly reach steady-state values and may be approximated by three constants.

The response of a simplified filter is shown in Figs. 8a through 8c. The process of design enters since it required a number of runs using different slopes and time constants to get a good match with the response of the complete filter. In fact, in the end, the same exponential gains were used for both po_s and po_e; the same exponential gains were used for β and γ. The total precomputed constants were 3 sway variable gains, 2 initial conditions for exponentials, and sets of the following 5 numbers which are changed at 10 discrete times: 2 time constants for exponentials (po_s and β) and 3 slopes for straight-line segments (α, d_x, d_y).

Implementation

A slight variation of this simplified filter was implemented in the Apollo Guidance Navigation and Control system. Part of the program was concerned with initialization for platform positions other than the one considered here. The optimum filter, once implemented, does not change for other platform positions; the measurements that the filter gets are made to simulate the standard platform configuration. For example, if the platform axes were vertical, north and east, and if the sign of the north accelerometer output were changed to simulate a south accelerometer, the filter output for the variables in the north direction need only be interpreted as negative of their true values. Similarly, for some platform positions it is necessary to resolve the measurements to simulate south and east accelerometers.

For some applications, it may be desirable to torque the south gyro at negative horizontal earth rate. The form of the filter and the filter gains do not change because perfect torquing is assumed; just add negative horizontal earth rate to the extrapolation of the angle β. It has also been found convenient to extrapolate the alignment angles according to simple first order ($\alpha=\alpha+\dot{\alpha}\,dt$, etc.). The sway variables are extrapolated according to a sway transition matrix whose elements can be changed to compensate for variations in launch vehicle parameters.

Once the optimum filter has been implemented according to the simple method outlined, it can be readily adapted to various problems of alignment and calibration. For example, consider the following system test procedure in which the platform axes are identified as x, y, z: 1) run a 10-min test with x up, y south, z east to determine y gyro bias drift; 2)at 10-min read out y drift and use the angle estimates to align the platform;continue to torque the platform at earth rate for 90 sec while counting pulses from the x accelerometer (vertical); 3) orient the platform to x down, y east, z south and run a 10-min test to determine z bias drift; 4) use the angle estimates at 10-min to align the platform and then torque for 90-sec at earth rate while counting x accelerometer pulses; and 5) torque the south gyro at horizontal earth rate for 45-min more while determining vertical drift (x gyro). The y bias as determined in 1 is d_z for this run.

This procedure takes about 68 min after which enough information is available to determine y gyro bias drift, z gyro bias, the sum of x gyro bias and acceleration sensitive drift, and x accelerometer bias and scale factor. One can readily imagine how an automated system test procedure can be set up to completely calibrate the system in the swaying spacecraft. The last step in the program would be an alignment run to ready the system for launch. And through all of this, the basic simplified optimum filter does not change.

Hardware Problems

The primary source of azimuth error is the uncertainty in east gyro drift which comes about from errors in the calibration of the gyro; the primary source of vertical drift error is due to variations in the east gyro drift during a test. If the east gyro has a large drift due to acceleration along its input axis, then it is desirable to keep the input axis almost horizontal during the test by torquing the south gyro at negative earth rate to minimize the variations in east gyro drift. Unfortunately, if the bias of the south gyro changes when it is torqued, then the south gyro calibration will yield two answers corresponding to the torqued and untorqued cases, respectively. (Torquing was done in the system test program.)

Another possible problem area appears when using pulsed integrating accelerometers. If either horizontal accelerometer has a large dead-zone for near zero inputs, large transients in the filter output will appear. The form of the transients will vary depending on the time during the test that the accelerometer goes through the dead-zone. Vertical drift estimation

is particularly sensitive to a dead-zone in the south accelerometer. In some cases a transient on the order of 800 meru has been observed; the filter never reached the correct value of drift at the end of 45 minutes because the vertical drift gain is small at the end of the end of the test. As a practical solution to the problem, the platform is deliberately offset (between steps 4 and 5 in the system test program) from the vertical before beginning a vertical drift test so that the accelerometer never goes through null. This problem is indicative of the strange results that can occur when the model for the system is incorrect. Philosophically, we have designed a total system test and as such it should indicate in some manner out of specification conditions which would then require lower testing. The model must therefore include all in-spec conditions and the designer's experience must be used to recognize out-of-spec situations.

Conclusion

The possibility exists for significant reduction in the errors in calibration systems by using statistical estimation techniques that make the most efficient use of all available information. It also appears likely that suitable simplifications can be made to adapt these techniques to practical computers as was illustrated with two examples.

Acknowledgement

This report was prepared under DSR Project 55-23870, sponsored by the Manned Spacecraft Center of the National Aeronautics and Space Administration through Contract NAS 9-4065 with the Instrumentation Laboratory, Massachusetts Institute of Technology, Cambridge, Mass. The publication of this report does not constitute approval by the National Aeronautics and Space Administration of the findings or the conclusions contained therein. It is published only for the exchange and stimulation of ideas.

References

1. Kalman, R.E., "A New Approach to Linear Filtering and Prediction Problems", Journal of Basic Engineering, Transactions of the ASME, Vol. 82, pp 35-45, March 1960.

2. Kalman, R.E. and Bucy, R.S., "New Results in Linear Filtering and Prediction Theory", Journal of Basic Engineering, Transactions of the ASME, Vol. 83, pp. 95-108, March 1961.

3. Potter, J.E., "A Guidance-Navigation Separation Theorem", Report RE-11, Massachusetts Institute of Technology, Experimental Astronomy Laboratory, Cambridge, Mass., August 1964.

4. Brock, L.D., "Application of Statistical Estimation to Navigation Systems", Ph.D. thesis, T-414, Department of Aeronautics and Astronautics, M.I.T., June 1965.

5. Schmidt, G.T., "Apollo System Test Group Memo No. 459", M.I.T. Instrumentation Laboratory, June 1965.

6. James, H.M., Nichols, N.B., and Phillips, R.S., Theory of Servomechanisms, p. 369, Dover Publications, New York, 1965.

7. Denhard, W.G., "Laboratory Testing of a Floated, Single-Degree-of-Freedom, Integrating, Inertial Gyro", Report R-105, MIT Instrumentation Laboratory, September 1956.

8. Schmidt, G.T. and Brock, L.D., "Statistical Estimation in Inertial Navigation Systems", AIAA Journal of Spacecraft and Rockets, Vol. 5, No. 2, pp. 146-154, February 1968.

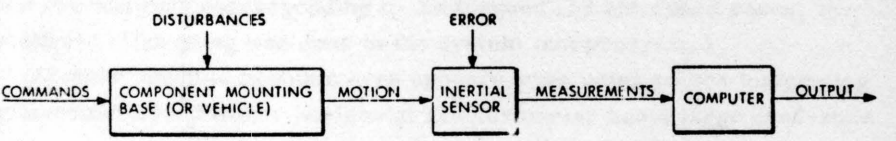


Fig. 1 Functional Form of Calibration System

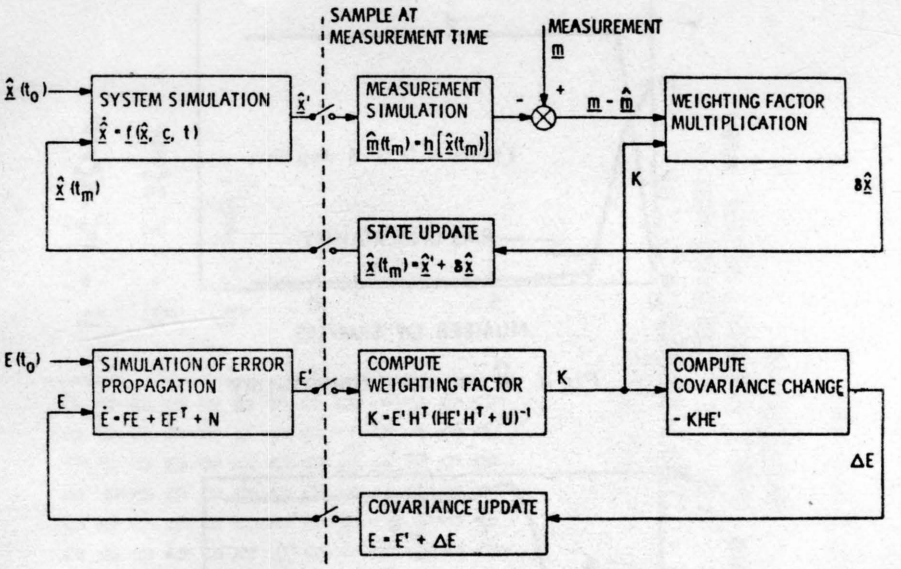


Fig. 2 Optimum Filter Functional Diagram

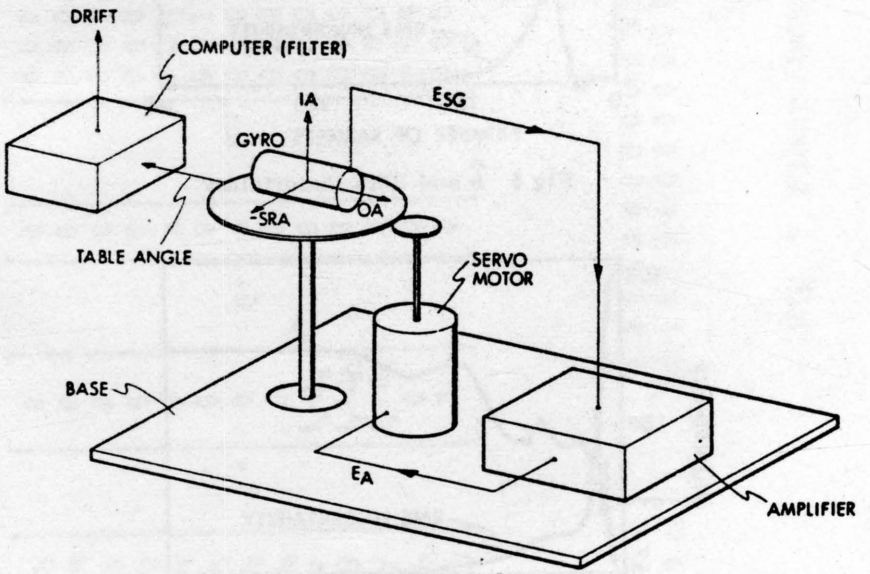


Fig. 3 Gyro Servo Test Calibration

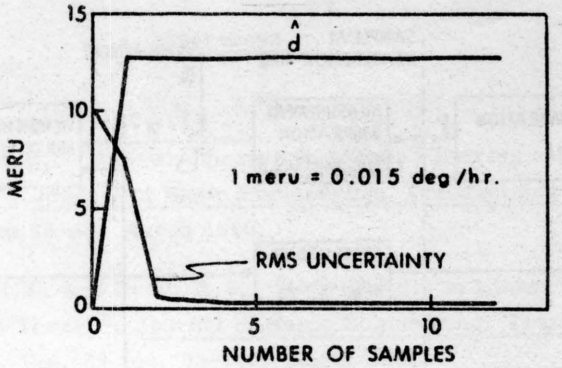


Fig 4 $\hat{d}$ and RMS Uncertainty

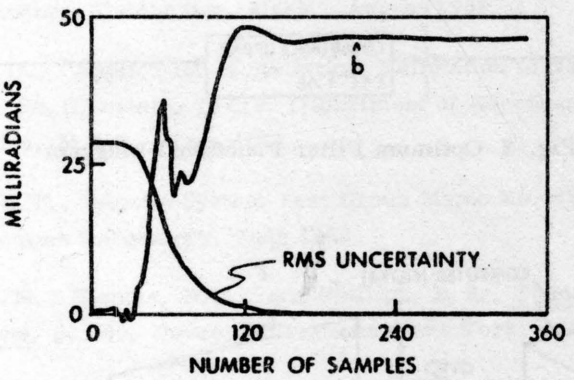


Fig 5 $\hat{b}$ and RMS Uncertainty

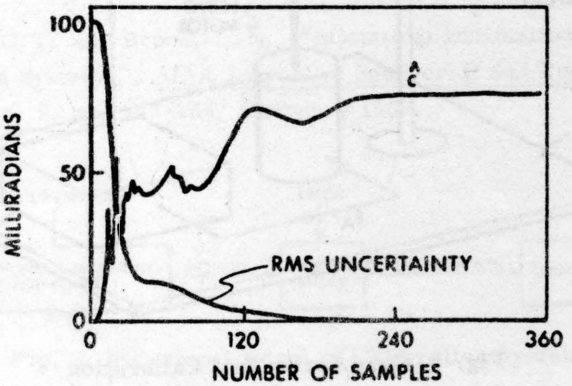


Fig 6 $\hat{c}$ and RMS Uncertainty

$$\underline{X} = \begin{bmatrix} p_{os} \\ p_{oe} \\ v_s \\ v_e \\ a_s \\ p_s \\ a_e \\ p_e \\ \alpha \\ \beta \\ \gamma \\ d_x \\ d_y \end{bmatrix} \quad \underline{c} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ t_x - \Omega_v \\ t_y + \Omega_h \\ t_z + d_z \\ 0 \\ 0 \end{bmatrix} \quad \underline{n} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ n \\ 0 \\ n \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$F = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\dot{g} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dot{g} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & f_1 & 0 & f_2 & f_3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & f_1 & 0 & 0 & f_2 & f_3 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \Omega_h & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \Omega_v & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\Omega_h & -\Omega_v & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$f_1 = -\omega_n^2 - 2\zeta\lambda\omega_n$$

$$f_2 = -\lambda - 2\zeta\omega_n$$

$$f_3 = -\lambda\omega_n^2$$

$$U = \begin{bmatrix} r_a & 0 \\ 0 & r_a \end{bmatrix} \quad H = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

N = 13x13 ZERO MATRIX EXCEPT FOR ROW 5, COLUMN 5 AND ROW 7, COLUMN 7 WHICH ARE BOTH nw

Fig. 7 Filter Matrices and Vectors

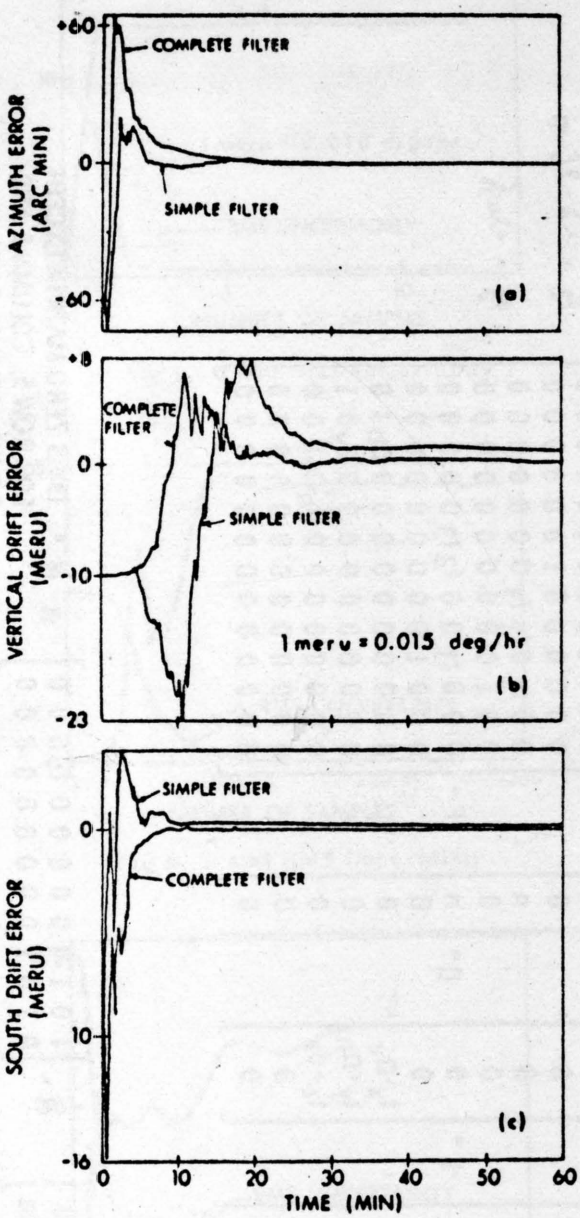


Fig. 8 Complete and Simplified Filter Performance

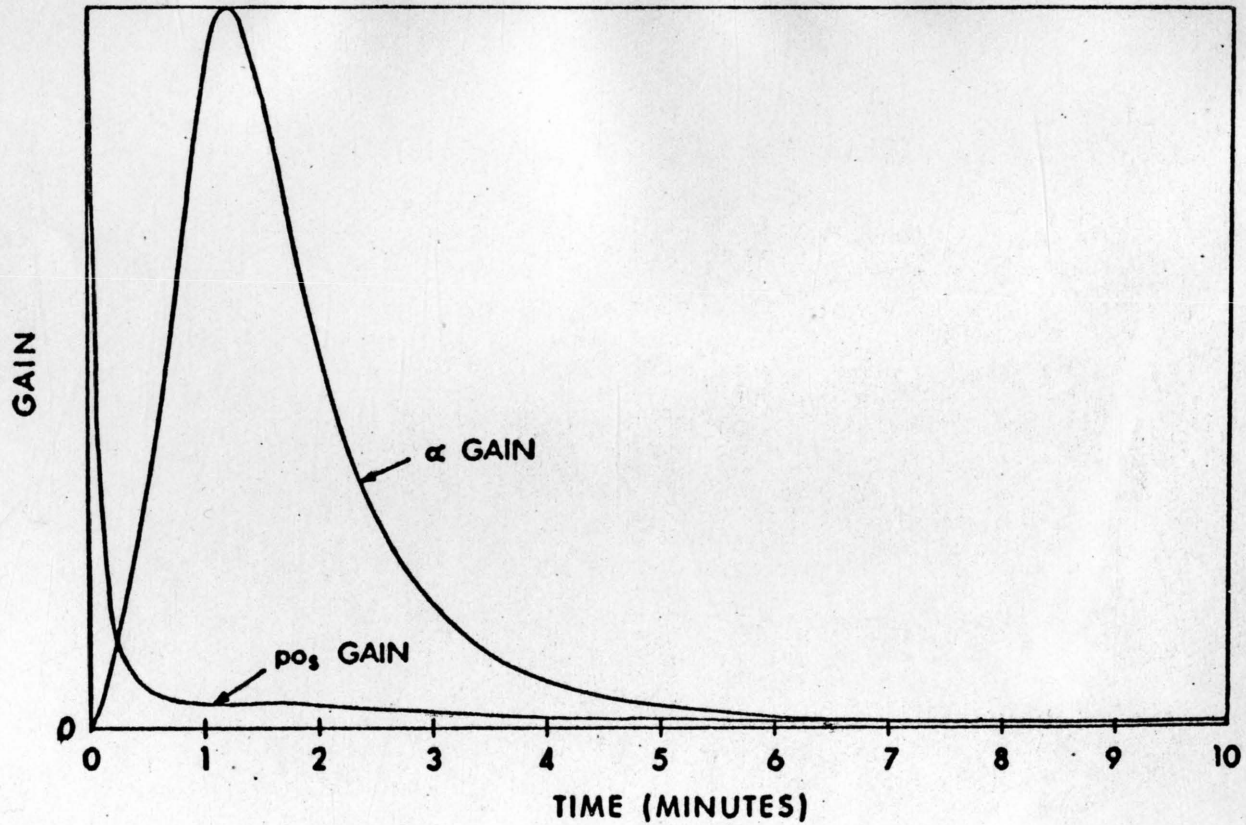


Fig. 9 Typical Gain Histories