"EFFECT OF MIXING ON REACTOR PERFORMANCE" AND "APPLICATION OF LINEAR PROGRAMMING TO CHEMICAL ENGINEERING"

by 884

Billy I Tsai

B. S., National Taiwan University, 1963

A MASTER'S REPORT

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Chemical Engineering

KANSAS STATE UNIVERSITY

Manhattan, Kansas

1968

Approved by:

Liang-tseng Fan

Major Professor

TABLE OF CONTENTS

EFFECT OF MIXING ON REACTOR PERFORMANCE

1. INTRODUCTION

The manner and extent to which flow behavior, i.e. mixing, affects the performance of continuous reactors have begun to be taken into consideration in recent years. While much progress has been made in understanding certain aspects of this problem, the task of developing a general treatment still appears to be formidable [18, 21].

The purpose of this report is to review some of the background information concerning the effect of mixing on the design of chemical reactors.

When a fluid flows through a process vessel, the condition of either plug flow or perfect mixing is often assumed in designing it. In practice, however, many systems do not conform to either of these assumptions; hence calculations based on them may be inaccurate. In order to describe actual flow systems which lie between the conditions of plug flow and perfect mixing, Danckwerts [2] introduced the concept of the residence time distribution (RTD). He further explained how the residence time distribution can be defined and measured for an actual system.

It has generally been assumed that the residence time distribution could well be used to determine the conversion of the chemical reaction, but this assumption is not always true. Kramers [9] compared two different reactor systems, both consisting of a tubular and a perfectly-stirred vessel in series, (Fig. 1). He found that they had identical

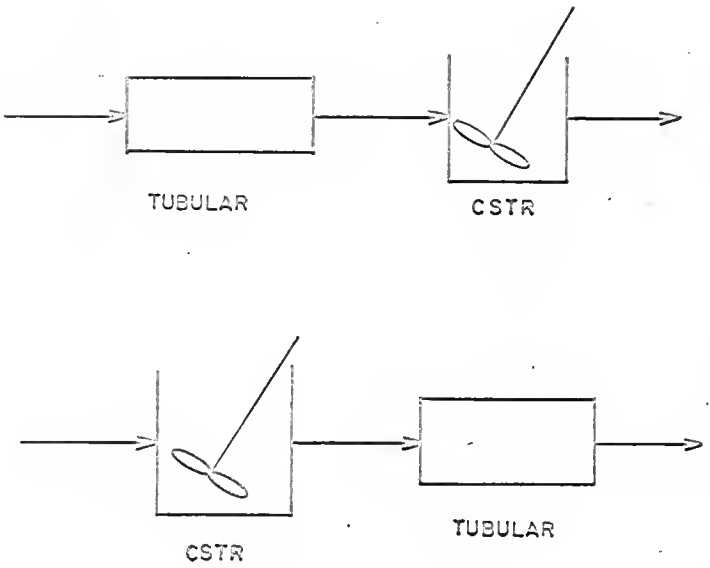

Fig. 1. Combined reactor systems.

residence time distributions but gave rise to different degrees of conversion for a chemical reaction of the second order.

Danckwerts [3] showed that the residence time distribution is sufficient to specify the degree of conversion only when the reaction is of the first order. Zwietering [20] showed later that the deviation in the degree of conversion between any two systems can be partially explained by the difference between the two age distributions. He showed further that with an arbitrary but known residence time distribution, it is possible to derive, for any reaction mechanism, limits between which the conversion must lie.

The work of these earlier investigators which is described above leads to the conclusions that, in order to give a fuller description of mixing effect on the performance of a chemical reactor, more information concerning mixing other than the knowledge of the residence time distribution is needed, and that it is convenient to be able to describe mixing within continuous flow systems in terms of two components, macromixing and micromixing [17].

The macromixing component specifies the variation in the residence times experienced by the molecules flowing through a flow system. The micromixing component, however, specifies the variation of environment experienced by the molecules during their passage through the system, [6, 7, 17, 18]. It will be made more evident later that the micromixing component is described in terms of, and is restricted by, the macromixing component. Macromixing is determined completely by the residence time distribution and will be discussed further in the next section.

2. MACROMIXING

Macromixing is the aspect of mixing accomplished by gross fluid motion in a flow system (chemical reactor). Specifically, macromixing describes the variations in holding times experienced by fluid elements passing through the reactor. A distribution function of holding times is called the residence time distribution (RTD) or the age distribution frequency of the exit stream from a vessel and it completely defines macromixing [2].

The residence time distribution of a system can be determined experimentally. Experimental techniques based on introducing a tracer into the inflowing stream and recording the resultant tracer concentration in the out-flowing stream have been summarized by Danckwerts [2].

The residence time distribution can be stated in a dimensionless form as $E(\theta)$, where $E(\theta)d\theta$ is the fraction of the fluid that has a dimensionless residence time between θ and $\theta + d\theta$. Here θ is equal to the dimensional residence time t divided by the mean residence time τ, in which τ is equal to the reactor volume divided by the volumetric flow rate.

The important mathematical properties of the residence time distribution function may be expressed in the following way:

1. $\int_0^\infty E(\theta)d\theta = 1, \qquad E(\theta) \geq 0 \quad \text{for } \theta \geq 0$

2. $\int_0^\infty \theta E(\theta)d\theta = 1.$

A function related to the residence time distribution and of special importance is the F function (or the cummulated age distribution function) which is defined as

$$F(\theta) = \int_0^\infty E(\theta)d\theta.$$

From the first property of the E function, it can be shown that

$$0 \leq F(\theta) \leq 1 \qquad \text{for } \theta \geq 0$$

$F(\theta)$ may be interpreted as the total fraction of the exiting stream at the steady state with a residence time less than or equal to θ.

Usually the two functions, E and F, are related to the step and impulse responses of a process and provide an overall description of the mixing inside the system relative to the direction of flow. In other words, they describe the macro scale mixing or simply macromixing in the system.

Concerning the residence time distribution (RTD), there are two extreme cases. One extreme case is the RTD of the plug flow reactor (PFR) in which all fluid elements spend the same length of time in residence while passing through the reactor. The RTD of a plug flow reactor is shown in Fig. 2. The other extreme is the RTD of a continuous stirred tank reactor (CSTR) in which all elements within the reactor have equal probability of leaving the reactor in the next dt units of time at any moment. The RTD of a CSTR is an exponential decay function as shown in Fig. 3. A non-ideal reactor can have any RTD falling between these two extremes. This is shown in Fig. 4 [7].

The above description of the residence time distribution suggests a

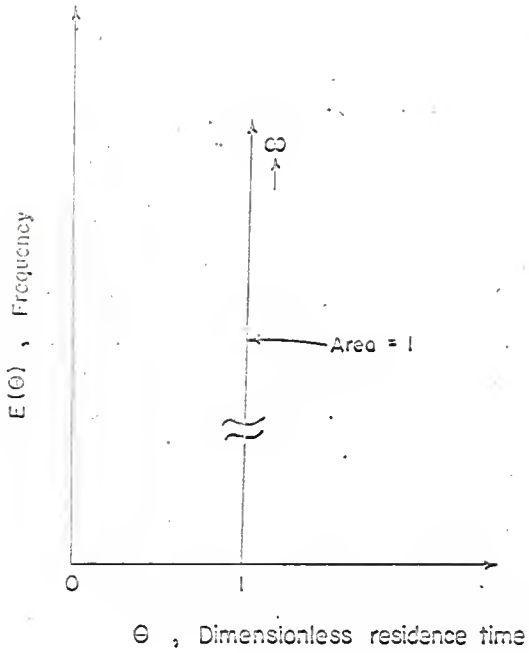

Fig. 2. Dimensionless residence time distribution of a plug flow reactor.

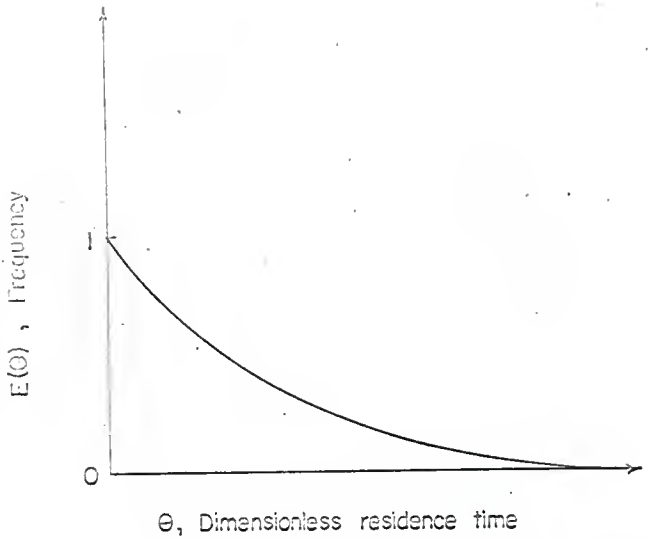

Fig. 3. Dimensionless residence time distribution of a continuous stirred reactor.

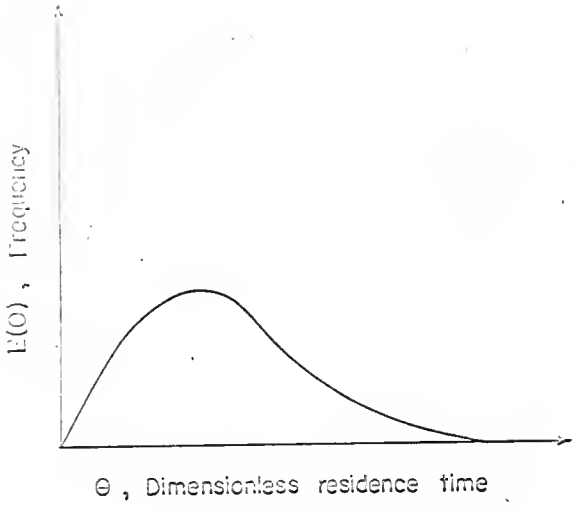

Fig. 4. Dimensionless residence time distribution of a non-ideal reactor.

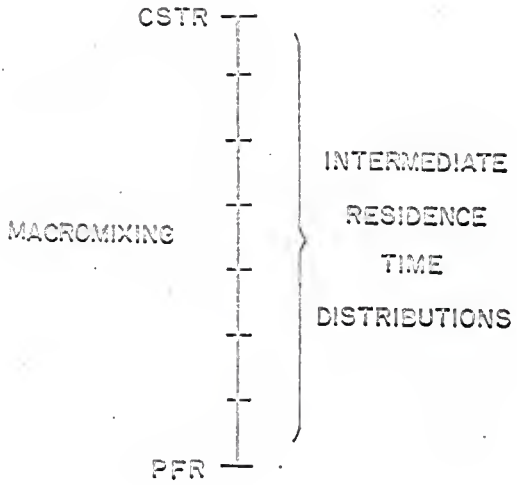

Fig. 5. Qualitative macromixing scale [17].

qualitative scale of macromixing bounded by the RTD of a plug flow reactor on one side and by the CSTR residence time distribution on the other limit. An arbitrary RTD may fall anywhere between these two limits. For conceptual purposes, a qualitative scale of macromixing may be visualized as shown in Fig. 5 [17].

However, in a process in which a chemical reaction occurs, a residence time distribution function alone, in general, will not be sufficient to predict the outlet conversion. This situation is due to the fact that a molecule may react before it passes out of the system. The possibility of reaction depends both on the molecule's duration of stay in the system and on its interaction with other molecules which it encounters during its stay in the system. In other words, to have a full description of conversion, both the macro scale mixing and micro scale mixing should be considered.

Although macromixing is only a partial description of mixing, it is sufficient to determine uniquely the conversion for the first order reaction [3, 5]. For reactions of an order other than the first, Zwietering [20] has shown that knowledge of the residence time distribution is sufficient to determine limits between which the conversion must lie. These limits correspond to two extreme states of micromixing, i.e., complete segregation and maximum mixedness, which will be discussed in the next section.

3. MICROMIXING [18]

Micromixing refers to all aspects of mixing not defined by RTD and is concerned with mixing on a molecular level. The concept of micromixing is perhaps best explained by considering fluid elements which are just entering or leaving a reactor [17, 18, 20]. An entering element of fluid is composed of molecules which are, in general, destined to have different residence times. All molecules in the entering element have life expectations which are equal to their individual residence times. Thus the environment of a molecule entering the reactor consists of molecules having identical ages but different life expectations. Similarly an element of fluid leaving the reactor is made up of molecules which have spent different lengths of time in residence. Since all molecules which have spent different lengths of time in residence. Since all molecules as they leave the reactor have zero life expectation, the molecules in the leaving element have ages which are equal to their individual residence times. The environment of a molecule leaving the reactor consists of molecules having an identical life expectation of zero but different ages. It is, therefore, clear that within the reactor a transition takes place from a grouping of molecules with identical ages to a grouping of molecules with identical life expectations. This transition of molecular groupings is called micromixing. These concepts of age and life expectation which provide an easy way of defining micromixing have been introduced by Zwietering [20].

The residence time distribution imposes certain restrictions on

micromixing in the following manner. As indicated by Zwietering [20], the residence time distribution uniquely determines the distribution of life expectations of the molecules within a reactor. It also determines the distribution of life expectations in the entering fluid. If these two types of life expectation distributions overlap, micromixing must take place to associate the entering molecules each having life expectation with the older molecules already in the reactor having the same life expectation. The degree to which these two distributions overlap is the measure of the micromixing which must take place within the reactor to satisfy the residence time distribution.

For the plug flow reactor, with its impulse residence time distribution, all the entering molecules have the same age of zero, and have life expectations equal to those of the residence times (the exit age). All of the molecules within the reactor have life expectations that are less than the residence times (the exit ages) because their ages are greater than zero. Thus plug flow permits no micromixing because no association of entering younger molecules with older molecules can take place. The continuous stirred tank reactor has a residence time distribution of the entering molecules, it is an exponential decay function. The life expectation distribution of the molecules within the reactor has the same exponential decay function. There is complete overlap of these two distributions and thus micromixing must occur.

Two natural extremes of micromixing can be explained according to how early or late the permissible association of entering molecules with older molecules already within the reactor occurs [20]. When the

association is as late as possible; i.e. at the exit of the reactor, the condition of segregated flow or (complete) segregation is to exist. When the association is as early as possible, the condition of maximum mixedness is said to exist.

Zwietering [20] devised models for the extreme cases of segregated flow and maximum mixedness. For the segregated flow, as shown in Fig. 6, the feed stream enters a plug flow reactor with side exits through which portions of the flow leave after satisfying their residence time requirements. The model classifies molecules within the reactor according to their age, . This arrangement mixes elements of the feed stream as late as possible; i.e., all of the micromixing takes place in the exit stream. For maximum mixedness, as shown in Fig. 7, the feed stream enters through side entrances to a plug flow reactor in order to satisfy the residence time distribution. The flow through each side entrance instantaneously mixes radially with the main flow. This model classifies molecules within the reactor according to their life expectation. This arrangement allows the mixing of the elements of the feed stream with elements of equal life expectation already within the reactor to take place as early as possible. In other words, the micromixing takes place as soon as the feed enters the reactor. More detailed discussion of Zwietering's model will be presented in section 5. As in the case of macromixing, a qualitative scale of micromixing may be envisioned for conceptual purposes as indicated in Fig. 8.

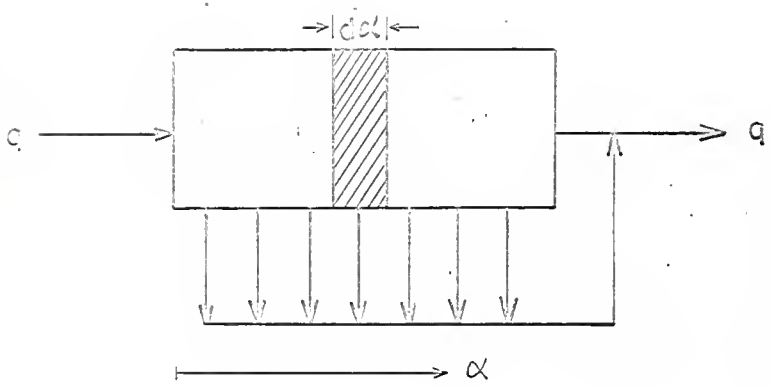

Fig. 6. Complete segregation model of Zwietering's plug flow reactor with side exits.

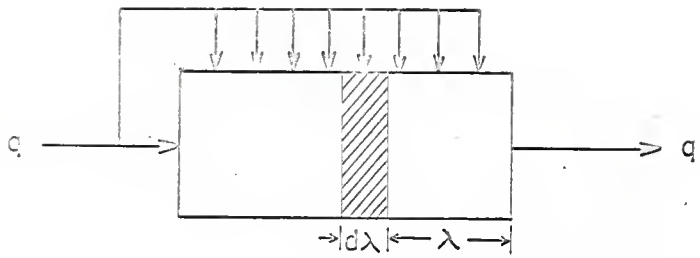

Fig. 7. Maximum mixedness model of Zwietering's plug flow reactor with side entrance.

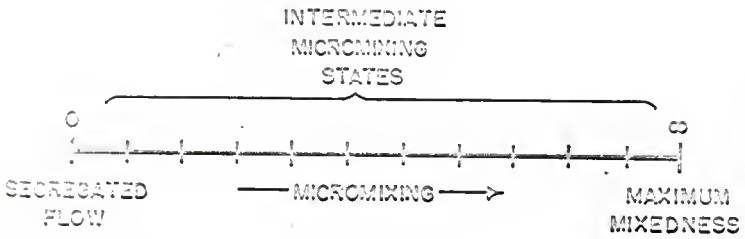

Fig. 8. Qualitative micromixing scale (17).

4. SEGREGATION

In section 3, it is stated that the two natural extremes of micromixing are, respectively, complete segregation and maximum mixedness. Some important aspects of segregated flow and maximum mixedness flow will be discussed more fully in sections 4 and 5, respectively.

Danckwerts [3] illustrated segregation by studying a "well stirred" continuous tank reactor in which an incoming element of fluid is uniformly dispersed throughout the reactor volume in a time much less than the mean residence time τ. Assuming that the residence time frequency distribution is $(1/\tau)\exp(-t/\tau)\,dt$, he showed that two limiting cases exist.

a. The fluid element entering the system is broken up into discrete fragments or streaks which are small in comparison with the volume of the system, and which are uniformly distributed throughout the system, and in which the molecules entering together remain together indefinitely.

b. The incoming fluid element is dispersed on a molecular scale in a time less than τ. In this state the mixture is chemically uniform and the neighborhood of any particular molecule does not tend to contain an excess of molecules which entered at the same time as itself.

Calculations of the overall reaction rate obtained in an isothermal reactor for the two limiting cases showed that the rates in the first cases would be higher for reaction orders greater than unity, lower for reaction orders less than unity and equal for reaction orders of unity,

if compared with those of the second case [3].

To provide a way of characterizing the mode of dispersion and thus to obtain a unique value of conversion, Danckwerts [3] introduced the concept of segregation. By considering the first case to be a completely segregated state, he defined a factor, called the degree of segregation, J, which varied from zero to unity as the state of the system varied from mixing on a molecular scale to complete segregation.

This degree of segregation, J, is defined as the variance of the mean ages of the points in the reactor divided by the variance of the ages of all the molecules of the system. One can write

$$J = \frac{\text{Var } \alpha_p}{\text{Var } \alpha} = \frac{\overline{(\alpha_p - \bar{\alpha})^2}}{\overline{(\alpha - \bar{\alpha})^2}} = \frac{\frac{1}{N}\sum_{p=1}^{N} (\alpha_p - \bar{\alpha})^2}{\frac{1}{M}\sum_{i=1}^{M} (\alpha_i - \bar{\alpha})^2} \tag{1}$$

where

α_i = age of molecule i, defined as the time which has elapsed since molecule i entered the system,

$\bar{\alpha}$ = mean age of all molecules which are at some particular moment in the system,

α_p = the mean age of the molecules at "point" p.

N = number of points considered in the analysis of the system.

M = number of molecules which are at some particular moment in the system.

One can see that J has a value of one for the plug flow reactor since anypoint contains molecules which have the same age, i.e. $\alpha = \alpha_p$.

In this case the variance of ages among points will be identical to the variance of ages among all the molecules. For any completely segregated flow, J will also have a value of one for the same reason.

In a continuous stirred reactor, where micromixing on a molecular scale may be attained instantaneously, the variance of ages among the points may be zero, i.e., $\sigma_p = \bar{\alpha}$, because all entering fluid will immediately be mixed on the molecular scale with all the fluid still remaining in the tank. In this case J is, therefore, zero. However, the flow may be completely segregated in this reactor, in which case J will be unity. It is, therefore, evident that a continuous stirred reactor may have any value of J from zero through one. This range represents the full scale of J.

For a system with a certain arbitrary residence time distribution which is neither that of a plug flow nor that of a continuous stirred tank reactor, the value of the upper limit of J is always one, because the concept of complete segregation is still applicable. A value of J equal to zero as the lower limit is, however, impossible because there is always a difference between the mean age of the molecules at a "point" and the mean age of all molecules in the system. It is necessary to designate a factor which is able to represent the lowest value attainable by J. The residence time distribution places this lower bound on J [20].

5. MAXIMUM MIXEDNESS AND ZWIETERING'S MODEL [17, 6]

Zwietering [20] generalized Danckwerts' concept of segregation [3] and introduced the concept of maximum mixedness which corresponds to the lowest value of J for a residence time distribution. The definition of the condition of maximum mixedness gives rise not only to the computation of a lower limit of J, but also to that of the minimum conversion for a reaction of an order higher than one (and the maximum conversion for an order of reaction lower than one).

For two series combinations of a plug flow and a continuous stirred tank reactor with the same RTD (see Fig. 1) it can be seen that mixing on a molecular scale throughout the whole system can never occur. At the same time it is possible to show that the conversions obtained from the two systems shown in Fig. 1 will, in general, be different even if the mode of dispersion in the stirred tank reactor is the same. If the reaction is isothermal and the order of reaction is greater than unity, the higher conversion will be obtained in the first configuration of Fig. 1 [6, 20]. Thus the first system may be considered to have a higher degree of segregation than the second has [6, 20].

Based on this example, Zwietering [20] provided an alternative definition of complete segregation. Recognizing that the only difference between the two reactor configurations shown in Fig. 1 is the fact that in the first configuration the mixing occurs at a comparatively later stage of the chemical reaction, he defined a completely segregated state to be one in which the mixing (or micromixing) of the fluid elements takes place as late as possible in the system, i.e., at the reactor

outlet. In contrast, in a system under the state of minimum segregation or the state of maximum mixedness, the mixing will take place as early as possible.

As briefly mentioned previously, Zwietering [20] used two simple diagrams to show the complementary nature of the states of complete segregation and maximum mixedness (see Figs. 6 and 7). Figure 6 shows a reactor configuration which corresponds to that of complete segregation. It consists of a plug flow reactor with a large number of extremely small side exits which are placed very close together. When the flow rate through the side exits is controlled to give a desired residence time distribution, the mixing of molecules with different ages will occur at the reactor outlet and the conditions required for complete segregation will be fulfilled.

A reactor corresponding to one with a state of maximum mixedness is depicted in Fig. 7. Since the only change from the first configuration (Fig. 6) is the reversion of the flow direction of the side streams giving rise to many entrances and only one exit, the residence time distribution will be unchanged. Since the plug flow reactor is assumed to have ideal or complete radial mixing, the entering molecules are mixed as soon as they enter the system. In addition, this model of the reactor satisfies the two conditions required for the state of maximum mixedness.

a. Molecules within a small group or a "point" have the same life expectations, i.e., molecules within a point in the system will leave at the same future moment.

b. Points having equal life expectations are mixed or at least have

identical age distributions.

It appears that the first condition is violated for the case of a continuous stirred (completely mixed) reactor or an ideal mixer. However, Zwietering showed that the residence time distribution for this system is a degenerate case and that the first condition is therefore not needed.

Specifically Zwietering [20] set up the models for the reactors with a completely segregated flow and a maximum mixedness flow in the following manner.

For segregated flow, he placed the entire flow in a plug flow reactor with side exits through which portions of the flow could leave after satisfying their residence time requirement (Fig. 6). The flow leaving between time α and $\alpha + d\alpha$ is $qE(\alpha)d\alpha$, where q is the flow rate. It is obvious that no backmixing is permitted in the plug flow reactor and that the immediate environment of each molecule consists of molecules with exactly the same history. A position in the plug reactor is measured from the inlet based on an age or α scale. The flow out of the side of the reactor between α and $\alpha + d\alpha$ has a fractional concentration of $r(\alpha) = \frac{C}{C_o}$ and a flow rate of $qE(\alpha)d\alpha$.

A material balance around a differential volume of the reactor between α and $\alpha + d\alpha$ provides the equation for the chemical conversion. The terms of the material balance of the reactant are as follow:

$$\text{mainstream inflow} = q\ \{1 - F(\alpha)\}\ r(\alpha)$$

$$\text{mainstream outflow} = q\ \{1 - F(\alpha + d\alpha)\}\ r(\alpha + d\alpha)$$

where in this case $F(\alpha)$ is the fraction of the total stream at position α

that has a residence time less than or equal to α, and $\{1 - F(\alpha)\}$ is the fraction of the total stream at position α that has a residence time greater than α.

outflow from the side exit = $r(\alpha)\, qf(\alpha)d\alpha$,

loss by reaction = $R(r)\, q\, \{1 - F(\alpha)\}\, d\alpha$.

A summary of the terms gives

$$0 = q\left\{\frac{\{1 - F(\alpha+d\alpha)\}\, r(\alpha+d\alpha) - \{1 - F(\alpha)\}\, r(\alpha)}{d\alpha}\right\} + qE(\alpha)r(\alpha)$$

$$+ qR(r)\, \{1-F(\alpha)\}$$

Since

$$\frac{dF(\alpha)}{d\alpha} = E(\alpha)$$

rearrangement and simplification of this expression gives the expected form

$$\frac{dr}{d\alpha} = - R(r) \qquad (2)$$

This formula is essentially the expression of the batch reaction rate. The outlet concentration is obtained by summing up contributions of the flows in the side exits. Thus the final concentration is

$$r_f = \int_0^\infty E(\alpha)\, r(\alpha)d\alpha \qquad (3)$$

or the final conversion is

$$x_f = \int_0^\infty E(\alpha)\, x_B(\alpha)\, d\alpha \qquad (4)$$

where

$$x_B(\alpha) = 1 - r(\alpha)$$

and

$$x_f = 1 - r_f .$$

This simply verifies the well-known formula of conversion at the exit x_f (or the exit concentration r_f) of a completely segregated reactor given below

$$x = \int_0^\infty E(t)\, x_B(t)dt, \qquad (5)$$

The derivation of the expression for the conversion for a reactor with a maximum mixedness flow may be carried out in the same way. But in this case the fluid enters through the side entrance at a coordinate point corresponding to its life expectation and all the fluid leaves together from the single exit of the reactor where $\lambda = 0$ (Fig. 7).

Note that a position in the plug flow reactor is measured from the exit based on a life expectation or λ scale. All the entering flows have the same inlet concentration. The flow entering between λ and $\lambda + d\lambda$ is $qE(\lambda)d\lambda$. The material balance of the reactant around the differential volume between λ and $\lambda + d\lambda$ is as follows:

$$\text{mainstream inflow} = q\ \{1 - F(\lambda + d\lambda)\}\ r_m(\lambda + d\lambda),$$

$$\text{side entrance inflow} = qE(\lambda)\ d\lambda(1),$$

$$\text{mainstream outflow} = q\ \{1 - F(\lambda)\}\ r_m(\lambda),$$

$$\text{loss by reaction} = R(r_m)\ \{1 - F(\lambda)\}\ q\ d\lambda.$$

A summing up of these terms gives

$$0 = q\ \left\{\frac{\left[1-F(\lambda+d\lambda)\right]\ r_m(\lambda+d\lambda) - \left[1-F(\lambda)\right]\ r_m(\lambda)}{d\lambda}\right\} + qE(\lambda)$$

$$- qR(r_m)\ \left[1-F(\lambda)\right]$$

and a rearrangement of the above expression gives

$$\frac{dr_m}{d\lambda} = R(r_m) - \frac{E(\lambda)}{1 - F(\lambda)}\left[1 - r_m(\lambda)\right], \qquad (6)$$

where $r_m(\lambda)$ is the fraction of the reactant which is unreacted at the point λ in a maximum mixedness reactor. The boundary condition for this equation is

$$\frac{dr_m}{d\lambda} = 0 \qquad \text{as} \qquad \lambda \to \infty \tag{7}$$

This equation can be integrated numerically and the value of $r_m(0)$ is the outlet concentration from which the conversion can be computed. The rigorous initial value of r_m is obtained by substituting the boundary condition in equation (6) and completing the limiting process of $\lambda \to \infty$ as shown below

$$0 = R\left[r_m(\infty)\right] - \left[1 - r_m(\infty)\right] \lim_{\lambda \to \infty} \frac{E(\lambda)}{1 - F(\lambda)} \tag{8}$$

Zwietering showed that when the limit of

$$\frac{E(\lambda)}{1 - F(\lambda)}$$

is known, $r_m(\infty)$ can be calculated. The value of $r_m(\infty)$ is insensitive to changes in λ with a magnitude of between three or four times the average residence time and ∞. Therefore, a numerical integration can be carried out by considering that $r_m = r_m(\infty)$ at $\lambda = 4\tau$ and starting

the integration there and completing it at $\lambda = 0$. This gives the concentration at the outlet of the reactor.

Zwietering [20] also indicated that a plug flow reactor is simultaneously in a state of complete segregation and a state of maximum mixedness. This condition occurs because this type of reactor has a residence time distribution which permits no freedom for different degrees of segregation; J is always equal to unity.

6. EFFECT OF MIXING ON CONVERSION

In previous sections, it has been indicated that each of the two components of mixing affects the chemical conversion in a different manner. While the effect of micromixing on the conversion may be discussed separately from that of macromixing, the latter must be discussed with reference to definite conditions of micromixing that have been assumed. In the following discussion of macromixing a condition of completely segregated flow will be assumed.

As mentioned previously, macromixing is represented by the distribution of the residence time about the mean. The fractional outlet conversion of a reactant in the case of a completely segregated flow can be expressed as

$$x_f = \int_0^\infty x_B(t)\ E(t)dt$$

where x_f is the final concentration of the product, $x_B(t)$ the concentration of the product in a batch reactor as a function of time and E(t) the residence time distribution. It can be shown mathematically [18] that the plug flow reactor and its corresponging impulse residence time distribution provide for the maximum chemical conversion for most types of chemical reactions.

As mentioned in section 2, the plug flow reactor and the continuous stirred tank reactor are normally taken to be the extreme cases so far as conventional reactor design considerations are concerned. Their residence time distributions, a delta function and an exponential decay, are considered as a limiting form of macromixing. It was also mentioned in section 3 that segregated flow and maximum mixedness flow are considered as two limiting cases of micromixing. We shall discuss the effects of these limiting cases of mixing on the conversion of the reactor.

Danckwerts [3] and Zwietering [20] showed that a unique prediction of the conversion could always be obtained for a plug flow reactor, and that the conversion obtained from a CSTR depends upon the degree of segregation as mentioned previously. Thus when a tubular reactor is compared with a CSTR, the degree of

segregation in the latter must be specified.

Examining the average reaction rates in the two reactors enabled Denbigh [4] provide a simple explanation of why, for all of the reaction orders that he investigated, a plug flow reactor always has a higher conversion than does a continuous stirred tank reactor with the state of maximum mixedness ($CSTR_{mm}$). For these systems in which the reaction rate decreases as the conversion increases, therefore, the conversion must increase along the reactor length in a tubular reactor, therefore, the average reaction rate will have a value between the high initial rate and the low final rate. In a $CSTR_{mm}$, however, the conversion at every point in the reactor is equal to the conversion at the reactor outlet, and the average rate is equal to the low value of the rate at the outlet. When the two systems are compared at the same value of outlet conversion, the tubular reactor will always have the higher average reaction rate, and therefore, will require a smaller reactor volume.

Danckwerts [3] also discussed the difference between the states of maximum mixedness and complete segregation in a CSTR in the following way. He took two portions of reaction mixture, which have different conversions and mixed them. Then he examined the average reaction rates in the resulting mixture under the two extreme states of micromising. It was shown that

the average rates were higher in the completely segregated state ($CSTR_{seg}$) when the reaction order was higher than unity, and lower when the reaction order was less than unity; and that the two states were equivalent when the reaction order was equal to unity. This result for first-order reactions might be expected. Since the kinetic equations for such reactions are linear, the super-position principle would apply, and the method of averaging would be immaterial. For second-order reactions, however, the kinetic equations are nonlinear and any dilution of the reactant concentrations will cause lower average reaction rates. The following example is given to illustrate major results discussed so far.

[Example 1] [5]

For an nth-order irreversible reaction in an isothermal plug flow reactor, it does not matter if a state of complete segregation of maximum mixedness exists. The material balance may be written as

$$\tau = \int_0^c \frac{dc}{-kc^n}$$

or

$$\tau = C_o \int_o^x \frac{dx}{KC^n} \tag{9}$$

or

$$C_0^{n-1} \tau = \int_o^x \frac{dx}{K(1-x)^n} \tag{10}$$

where τ is the space time, and x the fractional conversion,

$$\frac{C - C_0}{C_0} .$$

For a $CSTR_{mm}$, we obtain

$$\tau = \frac{C_o x}{-KC^n} \tag{11}$$

or

$$C_0^{n-1} \tau = \frac{x}{K(1-x)^n} \tag{12}$$

For a $CSTR_{seg}$, in addition to the rate equation, we also have to consider the segregation effect in order to determine the actual conversion. As mentioned in Section 5, the conversion in a completely segregated

reactor is given by Equation (5), i.e.

$$x_f = \int_o^\infty x_B(t)\ E(t)dt$$

where $x_B(t)$ is the conversion equation for an nth-order reaction in a batch reactor and is equivalent to the solution of

$$C_o^{n-1}\ t = \frac{1}{K(n-1)} \left\{\frac{1}{(1 - x_B(t)^{n-1}} - 1\right\}$$

The exit age distribution function for the CSTR, E(t) is equal to $\frac{1}{\tau}e^{-t/\tau}$.

The calculation of conversion with n = 1, 2, $\frac{1}{2}$ for the three preceeding cases were carried out by Douglas [5]. Equations of conversion for the plug flow tubular reactor, $CSTR_{seg}$, and $CSTR_{mm}$ are shown in Table 1. Results are shown in Figs. 9, 10, and 11. It is worth noting in Fig. 9 that for a first-order isothermal reaction, the $CSTR_{mm}$ and $CSTR_{seg}$ lead to the same conversion as pointed out by Danckwerts 3 . For a second-order reaction (Fig. 10) or a reaction of the order greater than one, the $CSTR_{seg}$ gives a higher conversion than $CSTR_{mm}$; however, for a one-half order reaction (Fig. 11) or a reaction of the order less than one, the $CSTR_{mm}$ yields a higher conversion that the $CSTR_{seg}$. For all of the above cases the conversion is only slightly affected by micromixing but is substantially affected by macromixing. The plug flow reactor always has the highest conversion.

The above examples are restricted to combinations of the extreme cases of both macromixing and micromixing. For non-extreme conditions of macromixing, or for any arbitrary RTD, one may also consider the different cases of micromixing, namely, segregated flow, maximum mixedness flow, and partially segregated flow.

For the first case, i.e. for any RTD with segregated flow, the conversion can be found by using Equation (5) as was done for the $CSTR_{seg}$:

$$x_f = \int_o^\infty x_B(t)\ E(t)dt.$$

For the second case, i.e. for any RTD with maximum mixedness, the chemical conversion can be found by using Equations (6) and (7).

$$\frac{dr_m}{d\lambda} = -\ R(r_m) - \frac{E(\lambda)}{1 - F(\lambda)}\ \{1 - r_m(\lambda)\}$$

and

$$\frac{dr_m}{d\lambda} = 0 \qquad \text{for} \quad \lambda \to \infty.$$

Table 1. Material Balance Equations [5]

PFR $$C_0^{n-1}\tau = \int_0^{\infty} \frac{dx}{K(1-x)^n}$$

$CSTR_{mm}$ $$C_0^{n-1}\tau = \frac{x}{K(1-x)^n}$$

$CSTR_{seg}$ $$x = \int_0^{\infty} x_B(t)\, E(t)\, dt$$

where $x_B(t)$ is equivalent to the solution of the plug flow tubular equation when τ is replaced by t.

For isothermal reactions:

	Tubular	$CSTR_{mm}$	$CSTR_{seg}$
n=1	$K = \ln \frac{1}{1-x}$	$K\tau = \frac{x}{1-x}$	$K\tau = \frac{x}{1-x}$
n=2	$KC_0\tau = \frac{x}{1-x}$	$KC_0\tau = \frac{x}{(1-x)^2}$	$x = 1 - \frac{\exp(\frac{1}{K} C_0\tau)}{KC_0\tau} E_i(\frac{-1}{KC_0\tau})$
n=3	$\frac{K\tau}{C_0^{1/2}} = 2(1-\sqrt{1-x})$	$\frac{K\tau}{C_0^{1/2}} = \frac{x}{\sqrt{1-x}}$	$x = \frac{K\tau}{C_0^{1/2}} - \frac{1}{2}\frac{K^2\tau^2}{C_0} + \frac{1}{2}\frac{K^2\tau^2}{C_0}\exp(-\frac{2C_0^{1/2}}{K\tau})$

$$0 \leq \frac{K\tau}{C_0^{1/2}} \leq 2$$

where E_i is the expontial integral, i.e.

$$E_i(y) = \int_y^{\infty} \frac{e^{-x}}{x}\, dx$$

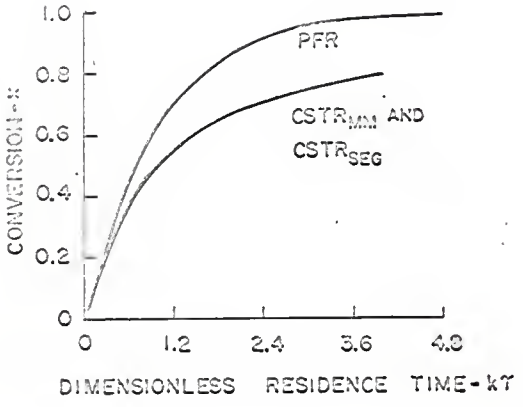

Fig.9. Conversion profiles, first-order isothermal reaction (5).

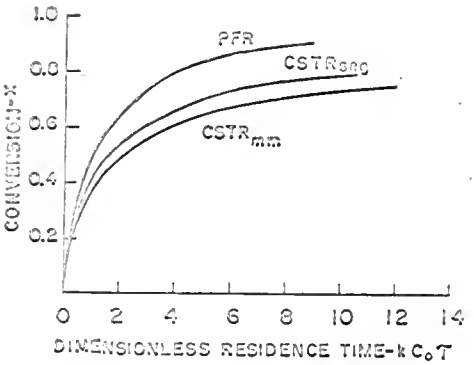

Fig.10. Conversion profiles second-order isothermal reaction (5).

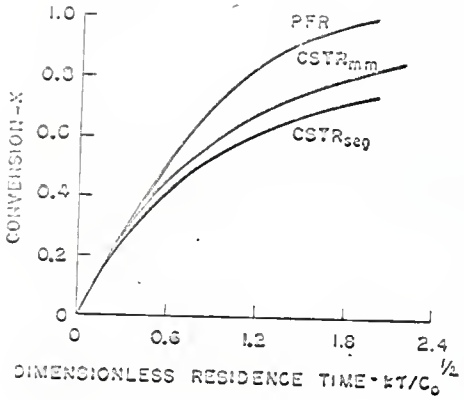

Fig.11. Conversion profiles, one-half-order, isothermal reaction (5).

For the third case, any RTD with partially segregated flow. there are various models of micromixing available for use. These will be discussed in the next section.

7. THE MODELING OF MICROMIXING

In Section 5, the extreme cases of micromixing, complete segregation and maximum mixedness, are discussed. Any condition lying between these two limits is called incomplete micromixing or partially segregated flow. The following is an extensive survey of the previous works concerning models of incomplete micromixing.

Weinstein and Adler [17, 18] proposed two simplified models of mixing, each of which requires only one micromixing parameter. The parameters in each of the two models are as follows:

(1) An age that separates the reactor into two parts - one part containing molecules younger than this age in a condition of complete segregation, the other part containing molecules older than this age in a condition of maximum mixedness. This picture leads to the use of the consecutive type model (see Fig. 12).

(2) An age that separates the entering fluid into two parts - one part containing molecules which will eventually have residence times of less than this age and always remain in a condition of complete segregation, and the other part containing molecules which will have residence times greater than this age and always remain in a condition of maximum mixedness. This picture leads to the parallel type model (see Fig. 13).

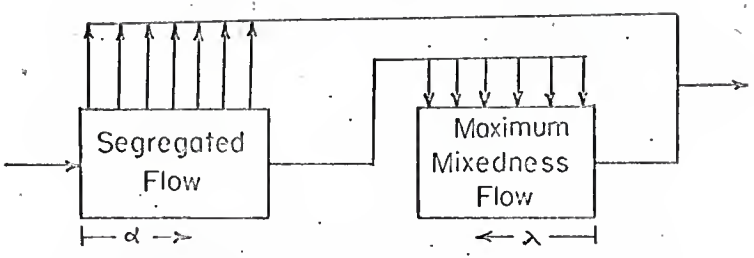

Fig. 12. Consecutive reaction model [18].

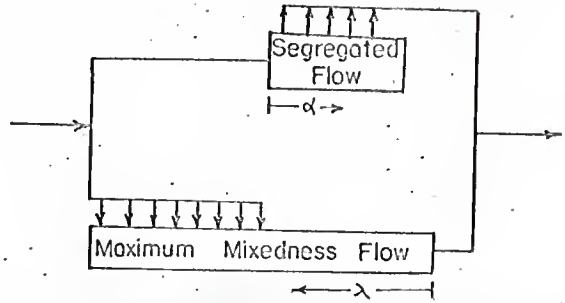

Fig. 13. Parallel reactor model [19].

The consecutive reactor model and the parallel reactor model are complementary in that each of them represents a different family of micromixing states. The consecutive reactor model is concerned with the age at which fluid begins to experience micromixing. The parallel reactor model is concerned with the fraction of the flow which experiences micromixing. It is felt that the consecutive reactor model better represents micromixing in most cases, since all of the fluid entering a system tends to remain segregated for at least a short time.

Rippin [14] established a quantitative relationship of the recycle reactor model (see Fig. 14) with the longitudinal diffusion model and the tanks in series model. He then showed that the recycle reactor can be used as a model of a flow reactor with incomplete mixing (macro). With regard to macromixing, the RTD of the recycle reactor becomes that of the plug flow reactor and that of the continuously stirred tank reactor as the recycle ratio R tends to zero and infinity respectively. An arbitrary RTD for a reactor which lies between that of a PFR and that of a CSTR may be obtained by proper choice of the recycle ratio R. With regards to micromixing, the recycle reactor itself is actually always at the state of maximum mixedness since it fulfills the condition of Zwietering's model of maximum mixedness, namely, that the association of entering molecules with older molecules is as early as possible (specifically at the inlet of the reactor). Thus, for this model, interpolation between the PFR and the CSTR is possible by changing the recycle ratio, but this change in the macromixing (i.e. the RTD) simultaneously causes the change in the condition of micromixing which may be characterized

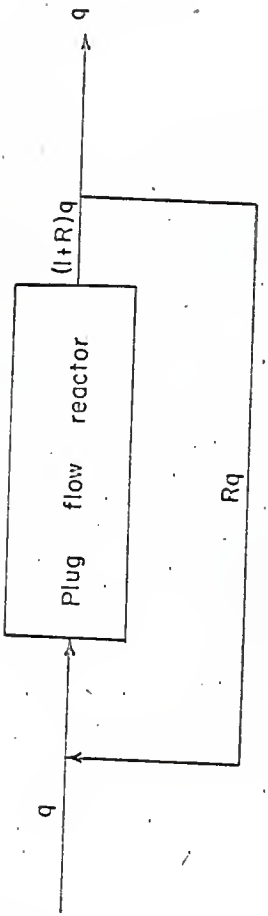

Fig. 14. Recycle reactor model [14].

by the degree of segregation J. No matter what the RTD of the recycle reactor may be, so far as micromixing is concerned the recycle reactor always is in the state of maximum mixedness, i.e. the state of minimum degree of segregation, corresponding to that RTD. This fact may be interpreted by Fig. 15. As the recycle ratio changes from zero to infinity, the condition of micromixing changes along the line of maximum mixedness, AC. If it is desired to study these two effects independently a micromixing model which can be used with any RTD is required. A single-parameter model of this type has been proposed by Ng and Rippen [12], and a multi-parameter model by Asbjornsen [1].

Ng and Rippin [12] recently proposed a two-environment model. By using this model, which allows the residence time distribution and the extent of micromixing in the system to be varied independently, the effect of micromixing is most clearly demonstrated. This procedure is not possible with some commonly used reactor models [14] such as the plug flow with diffusion, tanks in series or recycle reactor models. In this two-environment model the reactor is assumed to consist of an entering environment in which the fluid elements are completely segregated, and a leaving environment which is effectively a Zwietering maximum mixedness reactor. The rate of transfer between these two environments is determined by a transfer parameter R such that when R is zero there is no transfer from the entering environment to the leaving environment and the whole reactor is completely segregated. When R is infinitely large, the whole of the entering material is immediately transferred into a leaving environment such that the whole reactor

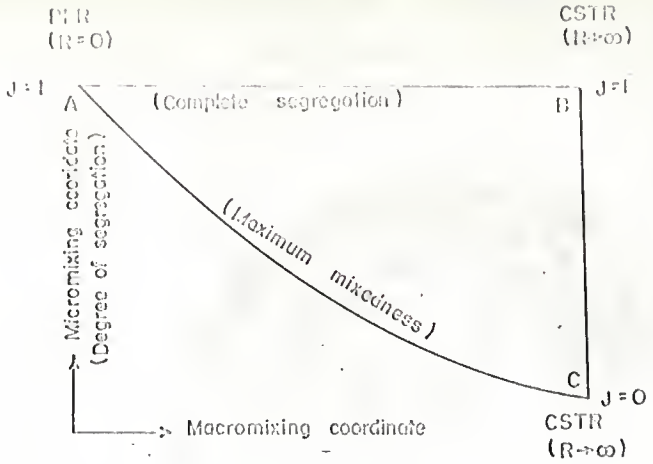

Fig. 15. Mixing conditions in the recycle reactor model.

becomes a maximum mixedness reactor. Intermediate values of R gives rise to different degrees of micromixing.

Rippin [15] developed a procedure for the evaluation of Danckwerts' degree of segregation for a two-environment model of a partially mixed (micro as well as macro) reactor from knowledge of the residence time distribution and the micromixing transfer parameter R. He subsequently applied this procedure to obtain an analytical relationship between the degree of segregation and the transfer parameter of the two-environment model for the residence time distributions equivalent to one and two well-stirred vessels. Rippin [15] also established the fact that different concentration histories may thus cause different conversion even in systems having the same residence time distribution and the same value of J. More specifically, only in the extreme cases of micromixing, i.e., maximum mixedness and complete segregation, will the degree of segregation accompanying the RTD provide the full information of conversion for any reaction mechanism. In the case of partial micromixing, the degree of segregation does not provide a good description of micromixing. Thus the degree of segregation together with RTD cannot uniquely determine the conversion. Therefore, both the concentration history and the RTD should be taken into consideration in determining the conversion for any chemical reactor of partial micromixing.

Asbjornsen [1] proposed a multi-parameter model of incomplete mixing obtained by using a fluid-flow network with a triangular structure and by subdividing the input and output of a system with a CSTR overall time

distribution into a number of segregated input and output streams as shown in Fig. 16. In this figure a circle indicates an ideal mixing stage and the number inside the circle denotes the mean residence time for this mixing stage. The network has two main input streams each divided into n parallel streams and these two main streams may be considered as two completely segregated inputs (segregated feed of reactants). The inter-mixing between the different streams in the system is simulated by ideal mixing stages at the nodes of the network. Asbjornsen also studied the effect of the degree of segregation on the overall conversion of the reactants in the following reaction system.

a. A second order, irreversible chemical reaction with equimolar feeds in the abiabatic reactor.

b. A second order, irreversible chemical reaction with equimolar feeds in the isothermal reactor.

c. An isothermal autocatalytic reaction with traces of the product in the input.

He further investigated the two extreme cases of micromixing completely segregated reactant feeds and thoroughly mixed reactant feeds for different types of flow distributions through the network.

Kattan [6] developed a mixing model capable of treating unmixed as well as premixed feed and of approximating a wide variety of mixing states in continuous flow chemical reactors. This model is based on random coalescense and redispersion occurring between fluid elements of equal life expectations. The model has been developed for homogeneous, isothermal systems but it can be used in simple dispersed systems. A stochastic

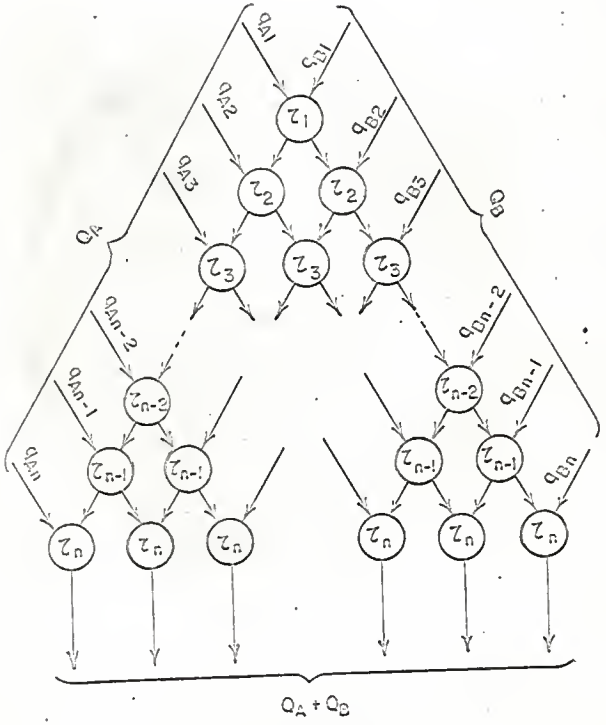

Fig. 16. A. Partly segregated flow network [1].

version based on the standard Monte Carlo technique is also implemented. Kattan [6] also studied three reactor problems. First, the effects of varying micromixing and residence time distributions were explored for several simple chemical reactions. Both premixed and unmixed feeds were treated. Secondly, the model was used to vary micromixing in an attempt to determine the yield of an arbitrary kinetic scheme involving side reactions. Finally, the model was used to simulate experimental results reported by others for a plug flow reactor.

Kattan and Adler [7] presented a conceptual framework for describing mixing in a continuous flow system based on the notions of the residence time distribution, residual life times (life expectations), and coalescense and redispersion of fluid elements. This framework is believed to be sufficiently comprehensive and flexible for use in any mixing situation. These investigators also derived a limited analytical formulation and compared it with three mixing models instituted by other investigators. They further suggested the Monte Carlo simulation as a practical means of implementing the model. Fukumi [16] studied the dynamics of an imperfect micromixing chemical reactor based on the diffusion model, the tanks in series model and the mixed model.

The models that have been mentioned above are all concerned with deterministic processes. There are mixing models of stochastic processes developed by Kattan and Adler [8] and Krambeck, Shinnar and Katz [10]. Kattan and Alder developed a simple stochastic mixing model for tubular reactors. The model is based on random collision and redispersions between elements of fluid. Krambeck, Shinnar and Katz modeled turbulent

chemical reactors by a network of stirred tanks with the stochastic nature of the mixing introduced by taking the interstage flows to be a stationary Markov process. They also discussed several general features of tracer experiments in these quasi-steady flows, together with the relations of the experimental result to the residence time distributions. They further analyzed the statistics of tracer experiments related on the one hand to the estimation of mixing parameters, and on the other to the forecasts of average yield from the reactor system under first order kinetics.

Makoto [11] proposed models of coalescence and redispersion for micromixing, and estimated the effects of micromixing on the conversion in continuous flow reactor using these models.

8. CONCLUSION

A single isothermal reaction is only slightly affected by micromixing, but is substantially affected by macromixing. In most cases a plug flow reactor which has the impulse type residence time distribution, gives rise to the maximum conversion of the reactant [5]. For an adiabatic reaction, the micromixing effects on conversion may be as great as that of the macromixing [5]. Sets of simultaneous isothermal reactions as well as reactions accompanied by large heat effects are greatly affected by both micromixing and macromixing [16]. Of course, a single or simultaneous isothermal reaction of the first order are completely governed by macromixing; micromixing has no effect.

The major significance of understanding the mixing effects on

reactor performance lies in the fact that they determine

(1) The condition of mixing most beneficial to reactor performance, and

(2) The sensitivity of reactor performance to changes in mixing condition.

The latter information is particularly valuable when a reactor is insensitive to micromixing. Then macromixing, conveniently defined by the residence time distribution, is established as the suitable criterion for design, scale-up and mathematical modeling.

PART II

APPLICATION OF LINEAR PROGRAMMING TO CHEMICAL ENGINEERING

APPLICATION OF LINEAR PROGRAMMING TO CHEMICAL ENGINEERING

Linear programming is a method of finding the maximum or minimum of a linear objective function subject to linear constraints. Most of the basic work was done in the 1940's by such men as F. L. Hitchcock, L. Kantorovitch, T. C. Koopmans, and G. B. Dantzig. In 1947, Dantzig formulated the general linear programming and developed the so-called simplex method of solution. Since then the applications of linear programming have been widespread and hundreds of technical articles and many books on linear programming have appeared. Many of the earlier uses of linear programming were made in the petroleum industry [4, 5, 16, 17]. Other extensive applications of linear programming including such problems as production scheduling, transportation, personnel, assignment, diet preparation, machine-loading, and materials-bending are presented in books by Llewellyn [11], Hadley [10], Gass [9], and Dantzig [6].

1. A SIMPLE EXAMPLE

Let us first consider the following simple problem as an illustration. Suppose that a function,

$$f = 12x_1 + 10x_2 \quad , \tag{1}$$

is to be maximized subject to the restrictions,

$$
\begin{aligned}
2x_1 + x_2 &\leq 8, \\
2x_1 + 3x_2 &\leq 12, \\
x_1 &\geq 0, \\
x_2 &\geq 0.
\end{aligned}
\tag{2}
$$

For this two dimensional case a graphical solution as shown in Fig. 1 is possible. The first restriction or constraint is represented by the area under the line,

$$x_2 = 8 - 2x_1, \tag{3}$$

and the second constraint by the area under the line,

$$x_2 = 4 - \frac{2}{3} x_1. \tag{4}$$

The third and fourth constraints jointly correspond to the area in the first quadrangle. It can be seen that we have to seek values of x_1 and x_2, which maximize the objective function in the shaded area.

Consider also a family of lines corresponding to the objective function, equation (1), which is a family of parallel lines with a slope of $-\frac{6}{5}$ and which moves away from the origin with increasing f. Maximizing f then is equivalent to seeking the line with the largest f containing a point (x_1, x_2), which belongs to the shaded region.

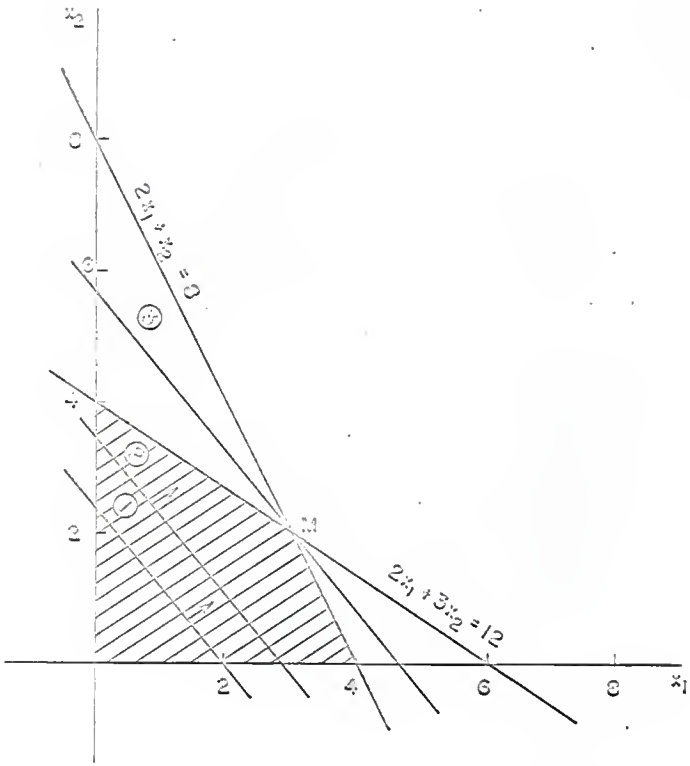

Fig. 1. Two dimensional maximun problem in linear programming.

Figure 1 shows such a line passing through point M. The values of

$$x_1 = 3 \qquad \text{and} \qquad x_2 = 2 \tag{5}$$

are obtained by the simultaneous solution of the two straight-line equations, equations (3) and (4). The corresponding maximum value of f is 56.

From this simple example we may draw the following conclusions.

(1) The optimal solution lies on the boundary of the shaded region or more specifically on the "corner".

(2) The shaded region is called a convex region, that is, all the points on any line segment connecting any two points of the feasible region.

This problem may also be stated in matrix-vector form as follows:

$$\max \begin{bmatrix} 12, & 10 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix},$$

subject to

$$\begin{bmatrix} 2 & 1 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \leq \begin{bmatrix} 8 \\ 12 \end{bmatrix},$$

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \geq \begin{bmatrix} 0 \\ 0 \end{bmatrix} ,$$

or more concisely as

$$\max \; c^T \underset{\sim}{x}$$

$$A\underset{\sim}{x} \leq \underset{\sim}{b} \qquad (6)$$

$$\underset{\sim}{x} \geq 0$$

where

$$\underset{\sim}{c} = \begin{bmatrix} 12 \\ 10 \end{bmatrix} , \quad \underset{\sim}{c}^T = \begin{bmatrix} 12 & 10 \end{bmatrix} , \quad A = \begin{bmatrix} 2 & 1 \\ 2 & 3 \end{bmatrix} ,$$

$$\underset{\sim}{b} = \begin{bmatrix} 8 \\ 12 \end{bmatrix}$$

2. THE PROBLEM OF LINEAR PROGRAMMING

A linear programming problem can now be generally stated as finding an n-dimensional vector, $(x_1, x_2, \ldots, x_n)$, or a set of values for n

variables, $\{x_1, x_2, \ldots, x_n\}$, which maximizes the linear objective or profit function,

$$c_1x_1 + c_2x_2 + \ldots + c_jx_j + \ldots + c_nx_n, \tag{7}$$

subject to m linear inequality constraints,

$$\begin{aligned}
&a_{11}x_1 + a_{12}x_2 + \ldots + a_{1j}x_j + \ldots + a_{1n}x_n \leq b_1,\\
&a_{21}x_1 + a_{22}x_2 + \ldots + a_{2j}x_j + \ldots + a_{2n}x_n \leq b_2,\\
&\qquad\qquad \ldots \qquad\qquad\qquad (8)\\
&a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{ij}x_j + \ldots + a_{in}x_n \leq b_i,\\
&\qquad\qquad \ldots\\
&a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mj}x_j + \ldots + a_{mn}x_n \leq b_m,
\end{aligned}$$

under the conditions that

$$x_j \geq 0, \qquad j = 1, 2, \ldots, n, \tag{9}$$

where a_{ij}, b_i and c_j are given constants and $b_i > 0$

The general linear programming problem may be stated in many forms. Some of the common ways of stating it are given below:

(1) Maximize

$$\sum_{j=1}^{n} c_j x_j$$

subject to

$$x_j \geq 0, \qquad j = 1, 2, \ldots, n$$

and

$$\sum_{j=1}^{n} a_{ij} x_j \leq b, \qquad i = 1, 2, \ldots, m.$$

(2) Maximize

$$\underset{\sim}{c}^T \underset{\sim}{x}$$

subject to

$$\underset{\sim}{x} \geq \underset{\sim}{0}$$

and

$$A\underset{\sim}{x} \leq \underset{\sim}{b}$$

where

$$\underset{\sim}{c}^T = (c_1, c_2, \ldots, c_n),$$

$$\underset{\sim}{x} = \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix}, \qquad A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdot & & & \\ \cdot & & & \\ \cdot & & & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

$$\underset{\sim}{b} = \begin{bmatrix} b_1 \\ b_2 \\ \cdot \\ \cdot \\ \cdot \\ b_m \end{bmatrix}, \qquad 0 = \begin{bmatrix} 0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{bmatrix}, \text{ (0 repeated n times).}$$

(3) Maximize

$$\underset{\sim}{c}^T \underset{\sim}{x}$$

subject to

$$A\underset{\sim}{x} \geq 0$$

and

$$x_1 \underset{\sim}{P}_1 + x_2 \underset{\sim}{P}_2 + \ldots + x_n \underset{\sim}{P}_n \leq \underset{\sim}{P}_0$$

where P_j, $j = 1, 2, \ldots, n$, is the jth column of the matrix, A, and $\underset{\sim}{P}_0 = \underset{\sim}{b}$. Note that without loss of generality, we can always consider a linear, programming problem as a problem of maximization since the minimum of an objective function is equivalent to the negative of the maximum of its negative, i.e.

$$\min (\underset{\sim}{c}^T \underset{\sim}{x}) = - \max (-\underset{\sim}{c}^T \underset{\sim}{x}).$$

3. THE DUAL PROBLEM

Corresponding to every linear programming problem there exists another linear programming problem called its dual. For example, the dual problem for the simple problem in Section 1, may be defined as

$$\min \underset{\sim}{b}^T \underset{\sim}{y} \tag{10}$$

$$\underset{\sim}{A}^T \underset{\sim}{y} \geq \underset{\sim}{c}$$

$$\underset{\sim}{y} \geq 0.$$

Or more specifically, it is used to minimize the new objective function, g, defined by

$$g = 8y_1 + 12y_2 \tag{11}$$

subject to the constraints

$$\begin{aligned} 2y_1 + 2y_2 &\geq 12 \\ y_1 + 3y_2 &\geq 10 \\ y_1 &\geq 0 \\ y_2 &\geq 0 . \end{aligned} \tag{12}$$

The procedure for solving this problem graphically is shown in Fig. 2. It is similar to that for solving the original maximization problem. Those values of y, which satisfy the constraints must lie in the shaded area bounded by the two straight lines,

$$\begin{aligned} y_2 &= 6 - y_1 \\ y_2 &= \frac{10}{3} - \frac{y_1}{3} . \end{aligned} \tag{13}$$

A family of lines representing the new objective function, g, moves decreasingly toward the origin in the region for positive g. Hence the smallest g which contains y_1 and y_2 in the shaded area passes through point $S(y_1 = 4, y_2 = 2)$ which corresponds to the simultaneous solution of the two straight line equations, equation (13). Thus, the minimum value of g is 56. We can immediately see that

$$\max f = \min g. \tag{14}$$

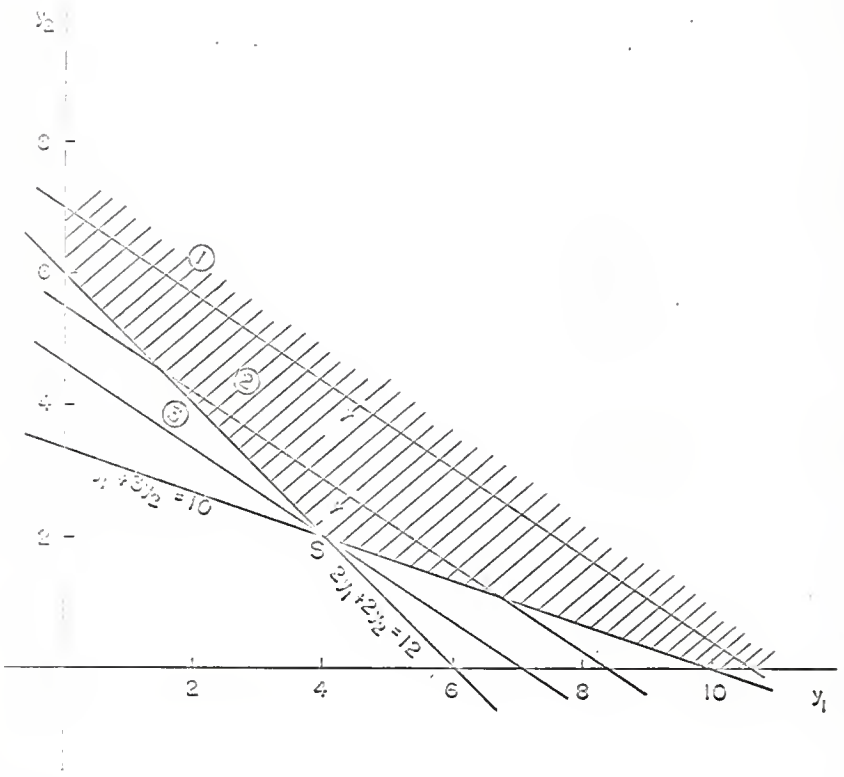

Fig. 2. Two dimensional minimum problem in linear programming.

The minimum point also lies on the boundary of a convex region of possible solutions.

Now in conjunction with the general linear programming problem presented in section 2, we can state its dual as finding an m-dimensional vector,

$$\begin{bmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ \cdot \\ y_m \end{bmatrix}$$

which minimizes the linear objective function,

$$g = b_1y_1 + b_2y_2 + \cdots + b_jy_j + \cdots + b_my_m, \tag{15}$$

subject to n linear inequality constraints,

$$\begin{aligned} a_{11}y_1 + a_{21}y_2 + \cdots + a_{j1}y_j + \cdots + a_{m1}y_m &\geq c_1, \\ a_{12}y_1 + a_{22}y_2 + \cdots + a_{j2}y_j + \cdots + a_{m2}y_m &\geq c_2, \\ a_{1i}y_1 + a_{2i}y_2 + \cdots + a_{ji}y_j + \cdots + a_{mi}y_m &\geq c_i, \\ a_{1n}y_1 + a_{2n}y_2 + \cdots + a_{jn}y_j + \cdots + a_{mn}y_m &\geq c_n, \end{aligned} \tag{16}$$

$$y_j \geq 0, \qquad j = 1, 2, 3, \ldots, m.$$

Or briefly in matrix notation, it can be stated as

$$\begin{aligned} &\min \underset{\sim}{b}^T \underset{\sim}{y} \\ &\underset{\sim}{A}^T \underset{\sim}{y} \geq c \\ &\underset{\sim}{y} \geq 0. \end{aligned} \tag{17}$$

The duality theorem [2, 5, 9] states that, for every linear programming problem, there is another, the dual problem, and that the solution of one exists if and only if the other has a solution. The maximum problem can be thought of as the dual of the minimum problem or the minimum problem can be thought of as the dual of the maximum problem. In either event, one problem is referred to as the primal or LP (linear program) and the other as the dual or DLP (dual linear program).

4. TRANSFORMATION OF THE PROBLEM

We shall first consider the LP and then turn our attention to the DLP. By introducing m new nonnegative variables

$$x_{m1}, x_{m2}, \ldots, x_{n+m}$$

which are called slack variables, we can replace the original inequality constraints,

$$
\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &\leq b_1, \\
a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &\leq b_2, \\
&\cdots \\
a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n &\leq b_m,
\end{aligned}
\tag{18}
$$

by the same number of equality constraints,

$$
\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \ldots + a_{1N}x_N &= b_1, \\
a_{21}x_1 + a_{22}x_2 + \ldots + a_{2N}x_N &= b_2, \\
&\cdots \\
a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mN}x_N &= b_m,
\end{aligned}
\tag{19}
$$

where

$$x_j \geq 0, \qquad j = 1, 2, 3, \ldots, N,$$

$$N = n + m,$$

$$a_{1,n+1} = a_{2,n+2} = a_{3,n+3} = \ldots = a_{mN} = 1,$$

and all other a_{ij} which have been newly introduced in this step are zero. The objective function, f, may now be written as

$$f = c_1x_1 + c_2x_2 + \ldots + c_Nx_N \tag{20}$$

where

$$c_{n+1} = c_{n+2} = \ldots = c_N = 0.$$

In matrix notation, these transformed equations can be written as

$$\begin{bmatrix} a_{11} & a_{12} \cdots a_{1n} & 1 & 0 \cdots 0 \\ a_{21} & a_{22} \cdots a_{2n} & 0 & 1 \cdots 0 \\ \cdot & & \cdot & \cdot \\ \cdot & & \cdot & \cdot \\ \cdot & & \cdot & \cdot \\ a_{m1} & a_{m2} \cdots a_{mn} & 0 & 0 \cdots 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_N \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \cdot \\ \cdot \\ \cdot \\ b_m \end{bmatrix}, \tag{21}$$

$$f = \begin{bmatrix} c_1 & c_2 \cdots c_N \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_N \end{bmatrix} \tag{22}$$

By defining

$$\underset{\sim}{P}_j = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \cdot \\ \cdot \\ \cdot \\ a_{mj} \end{bmatrix}, \qquad \underset{\sim}{P}_0 = \begin{bmatrix} b_1 \\ b_2 \\ \cdot \\ \cdot \\ \cdot \\ b_m \end{bmatrix}$$

$$\underset{\sim}{x} = \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_N \end{bmatrix}, \qquad \underset{\sim}{x}^T = \begin{bmatrix} x_1 & x_2, & \ldots, & x_N \end{bmatrix},$$

$$\underset{\sim}{c} = \begin{bmatrix} c_1 \\ c_2 \\ \cdot \\ \cdot \\ \cdot \\ c_N \end{bmatrix}, \qquad \underset{\sim}{c}^T = \begin{bmatrix} c & c & , \ldots, & c \end{bmatrix},$$

The linear programming problem or simply linear program may be written as

$$\max \underset{\sim}{c}^T \underset{\sim}{x} , \qquad \underset{\sim}{x} \geq 0$$

subject to

$$P_0 = \sum_{j=1}^{N} x_j P_j \quad . \tag{23}$$

5. SIMULTANEOUS LINEAR EQUATIONS AND GAUSS-JORDON REDUCTION

A system with m simultaneous linear equations with n unknown can be written as,

$$\begin{aligned}
&a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \ , \\
&a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \ , \\
&\quad \vdots \qquad\qquad\qquad\qquad \vdots \\
&\quad \vdots \qquad\qquad\qquad\qquad \vdots \\
&\quad \vdots \qquad\qquad\qquad\qquad \vdots \\
&a_{m1}x_1 + a\ \ x_{m2}{}^{+}{}_2 \ldots + a_{mn}x_n = b_m \ .
\end{aligned} \tag{24}$$

This equation can also be written in short form as

$$\underset{\sim}{A}\underset{\sim}{x} = \underset{\sim}{b} \ . \tag{25}$$

The solutions of underdetermined and determined sets of equations can be obtained by means of the Gauss-Jordon reduction [2, 10]. This method is a specialization of the Gauss' method for inversion of non-singular square matrices by elementary matric operation. If $a_{11} \neq 0$, dividing the first equation of equation (24) by a_{11}, we obtain

$$x_1 + \frac{a_{12}}{a_{11}} x_2 + \frac{a_{13}}{a_{11}} x_3 + \ldots + \frac{a_{1n}}{a_{11}} x_n = \frac{b_1}{a_{11}}$$

or

$$x_1 + b_{12}x_2 + b_{13}x_3 + \ldots + b_{1n}x_n = c_1$$

Multiply this equation by a_{21} and subtracting the result from the second equation, we obtain a new equation with the variable x_1 eliminated as

$$b_{22}x_2 + b_{23}x_3 + \ldots + b_{2n}x_n = c_2$$

where

$$b_{22} = a_{22} - a_{21}\frac{a_{12}}{a_{11}} = a_{22} - a_{21}b_{12}$$

$$b_{23} = a_{23} - a_{21}\frac{a_{13}}{a_{11}} = a_{23} - a_{21}b_{13}$$

$$\vdots \qquad \vdots \qquad \vdots \qquad \vdots \qquad \vdots$$

$$b_{2n} = a_{2n} - a_{21}\frac{a_{1n}}{a_{11}} = a_{2n} - a_{21}b_{1n}$$

$$c_2 = b_2 - a_{21}\frac{b_1}{a_{11}} = b_2 - a_{21}c_1$$

Repeatedly using the same procedure in the third, fourth, ..., nth equations obtain

$$\begin{aligned}
&x_1 + b_{12}x_2 + b_{13}x_3 + \ldots + b_{1n}x_n = c_1,\\
&\quad b_{22}x_2 + b_{23}x_3 + \ldots + b_{2n}x_n = c_2,\\
&\quad b_{32}x_2 + b_{33}x_3 + \ldots + b_{3n}x_n = c_3,\\
&\quad \cdot \qquad\qquad\qquad\quad \cdot \quad \cdot\\
&\quad \cdot \qquad\qquad\qquad\quad \cdot \quad \cdot\\
&\quad \cdot \qquad\qquad\qquad\quad \cdot \quad \cdot\\
&\quad b_{m2}x_2 + b_{m3}x_3 + \ldots + b_{mn}x_n = c_m.
\end{aligned} \tag{26}$$

If $b_{22} \neq 0$, we can divide the second of these equations by b_{22} and multiply by b_{12} and subtract the resulting equation from the first equation. This procedure eliminates x_2 from the first equation and the same procedure may be carried out to eliminate x_2 from each of the other equations. A new system of equations then has the form

$$
\begin{aligned}
x_1 + \quad & + c_{13}x_3 + c_{14}x_4 + \ldots + c_{1n}x_n = d_1, \\
x_2 + \; & c_{23}x_3 + c_{24}x_4 + \ldots + c_{2n}x_n = d_2, \\
& c_{33}x_3 + c_{34}x_4 + \ldots + c_{3n}x_n = d_3, \\
& \cdot \qquad\qquad\qquad \cdot \quad \cdot \\
& \cdot \qquad\qquad\qquad \cdot \quad \cdot \\
& \cdot \qquad\qquad\qquad \cdot \quad \cdot \\
& c_{m3}x_3 + c_{m4}x_4 + \ldots + c_{mn}x_n = d_m.
\end{aligned}
\tag{27}
$$

If we continue in this way, we shall arrive at a set of equations having the following form.

$$
\begin{aligned}
x_1 + \alpha_{1,r+1}x_{r+1} + \alpha_{1,r+2}x_{r+2} + \ldots + \alpha_{1n}x_n &= \beta_1, \\
x_2 + \alpha_{2,r+1}x_{r+1} + \alpha_{2,r+2}x_{r+2} + \ldots + \alpha_{2n}x_n &= \beta_2, \\
\cdot \qquad\qquad\qquad\qquad\qquad\qquad & \quad \cdot \\
\cdot \qquad\qquad\qquad\qquad\qquad\qquad & \quad \cdot \\
\cdot \qquad\qquad\qquad\qquad\qquad\qquad & \quad \cdot \\
x_r + \alpha_{r,r+1}x_{r+1} + \alpha_{r,r+2}x_{r+2} + \ldots + \alpha_{rn}x_n &= \beta_r, \\
0 &= \beta_{r+1}, \\
0 &= \beta_{r+2}, \\
& \cdot \\
& \cdot \\
& \cdot \\
0 &= \beta_m .
\end{aligned}
\tag{28}
$$

The number r is called the "rank" of the matrix A. Note that $r \leq \min(m,n)$. If $r < n$, equation (24) is called underdetermined; $r = n$, determined. We shall discuss all possible cases below

i) $m < n$

If $r < m$, the only way there can be solution to this set of equations is for $\beta_{r+1} = 0$, $\beta_{r+2} = 0$, ..., $\beta_m = 0$. The solution can then be written as

$$
\begin{aligned}
x_1 &= \beta_1 - \alpha_{1,r+1}x_{r+1} - \alpha_{1,r+2}x_{r+2} - \cdots - \alpha_{1n}x_n \,, \\
x_2 &= \beta_2 - \alpha_{2,r+1}x_{r+1} - \alpha_{2,r+2}x_{r+2} - \cdots - \alpha_{2n}x_n \,, \\
&\;\;\vdots \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \vdots \\
x_r &= \beta_r - \alpha_{r,r+1}x_{r+1} - \alpha_{r,r+2}x_{r+2} - \cdots - \alpha_{rn}x_n \,,
\end{aligned}
\tag{29}
$$

where if we set arbitrary values for x_{r+1}, x_{r+2}, ..., x_n, the values of x_1, x_2, ... x_r are uniquely determined. Here x_1, x_2, ..., x_n are called dependent variables or basic variables and x_{r+1}, ..., x_n are called independent variables or non-basic variables. The special solution obtained by setting the independent variables equal to zero and solving for the dependent variables is called a basic solution. Thus the basic solution of equation (24) is the following:

$$x_1 = \beta_1, \quad x_2 = \beta_2, \ \ldots \ x_r = \beta_r ,$$

$$x_{r+1} = x_{r+2} = \ldots = x_n = 0 .$$

A basic solution is degenerate if the values of one or of more of the dependent (basic) variables are zero. In particular, the basic solution of equation (24) is degenerate if $\beta_i = 0$, $i = 1, 2, \ldots, r$ for at least one i. If $r = m$, the values of $x_{m+1}, x_{m+2}, \ldots, x_n$ may be arbitrarily set and $x_1, x_2, \ldots, x_m$ are uniquely determined. If we set the values of the (n-m) variables $x_{m+1}, x_{m+2}, \ldots, x_n$ equal to zero, then the solution to the resulting system of equations is called a basic solution. The m variables which can be different from zero are called basic variables.

ii) $m > n$

There are more equations than unknowns. Since $r \leq n < m$, in order to have consistent solutions one must have $\beta_{r+1} = 0$, $\beta_{r+2} = 0, \ldots, \beta_m = 0$ and the last $m - r$ equations provide no useful information as far as the solution is concerned.

iii) $m = n$

There are an equal number of equations and unknowns. If $r < n$, then $(n-r)$ equations are dependent and can be ignored. A solution will exist only when $\beta_{r+1} = 0, \ldots, \beta_n = 0$. If $r = n$, a unique solution exists.

The above discussion may be summarized very neatly in terms of the ranks of two matrices. The first matrix is

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ . & . & & . \\ . & . & & . \\ . & . & & . \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \tag{30}$$

and the second matrix, called the argument matrix of A, is

$$\text{Aug } A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ . & . & & . & . \\ . & . & & . & . \\ . & . & & . & . \\ a_{m1} & a_{m2} & & a_{mn} & b_m \end{bmatrix} . \tag{31}$$

Carrying out the Gauss-Jordon reduction by pre-multiplying or post-multiplying A and Aug A respectively by suitable matrices, we get B and Aug B in the form

$$
B = \begin{bmatrix}
1 & 0 & 0 & \dots & 0 & \alpha_{1,r+1} & \alpha_{1,r+2} & \cdots & \alpha_{1n} \\
0 & 1 & 0 & \dots & 0 & \alpha_{2,r+1} & \alpha_{2,r+2} & \cdots & \alpha_{2n} \\
0 & 0 & 1 & \dots & 0 & \alpha_{3,r+1} & \alpha_{3,r+2} & \cdots & \alpha_{3n} \\
\cdot & \cdot & \cdot & & \cdot & \cdot & \cdot & & \cdot \\
\cdot & \cdot & \cdot & & \cdot & \cdot & \cdot & & \cdot \\
\cdot & \cdot & \cdot & & \cdot & \cdot & \cdot & & \cdot \\
0 & 0 & 0 & \dots & 1 & \alpha_{r,r+1} & \alpha_{r,r+2} & \cdots & \alpha_{rn} \\
0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\
0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\
\cdot & \cdot & \cdot & & \cdot & \cdot & \cdot & & \cdot \\
\cdot & \cdot & \cdot & & \cdot & \cdot & \cdot & & \cdot \\
\cdot & \cdot & \cdot & & \cdot & \cdot & \cdot & & \cdot \\
0 & 0 & 0 & \dots & 0 & 0 & 0 & & 0
\end{bmatrix} \tag{32}
$$

and

$$\text{Aug B} = \begin{bmatrix} 1 & 0 & 0 \ldots 0 & \alpha_{1,r+1} & \alpha_{1,r+2} & \cdots & \alpha_{1n} & \beta_1 \\ 0 & 1 & 0 \ldots 0 & \alpha_{2,r+1} & \alpha_{2,r+2} & \cdots & \alpha_{2n} & \beta_2 \\ 0 & 0 & 1 \ldots 0 & \alpha_{3,r+1} & \alpha_{3,r+2} & \cdots & \alpha_{3n} & \beta_3 \\ \cdot & \cdot & \cdot \quad \cdot & \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & \cdot \quad \cdot & \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & \cdot \quad \cdot & \cdot & \cdot & & \cdot & \cdot \\ 0 & 0 & 0 \ldots 1 & \alpha_{r,r+1} & \alpha_{r,r+2} & \cdots & \alpha_{rn} & \beta_r \\ 0 & 0 & 0 \ldots 0 & 0 & 0 & \ldots & 0 & \beta_{r+1} \\ 0 & 0 & 0 \ldots 0 & 0 & 0 & \ldots & 0 & \beta_{r+2} \\ \cdot & \cdot & \cdot \quad \cdot & \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & \cdot \quad \cdot & \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & \cdot \quad \cdot & \cdot & \cdot & & \cdot & \cdot \\ 0 & 0 & 0 \ldots 0 & 0 & 0 & \ldots & 0 & \beta_m \end{bmatrix} \tag{33}$$

If $\beta_{r+1} = 0$, $\beta_{r+2} = 0$, ... $\beta_m = 0$, equation (26) will have a solution and it follows necessarily that rank B is equal to rank aug B and rank A is equal to rank Aug A. Conversely, if rank A is equal to rank Aug A, the set of equations will have a solution. Summing up, we get an important theorem for linear simultaneous algebric equations. The necessary and sufficient conditions for a set of linear simultaneous equations to have a solution is that the rank of the matrix of the coefficients must be the same as the rank of the augment matrix of the coefficients [2].

A second important theorem deducted from the above discussion may be stated as follows. If the rank of the matrix of the coefficients in

a set of linear simultaneous algebraic equations is r and is the same as the rank of the argument matrix, and if n is the number of unknowns, the values of (n-r) of the unknowns may be arbitrarily assigned and the remaining r unknowns are uniquely determined, provided that the matrix of the coefficients of the remaining r unknowns has rank r [2].

Let us now return to considering the reduced system [see equation (28)] in the soluble case of the nonhomogeneous system equations $A\underset{\sim}{x} = \underset{\sim}{b}$. If r is the common value of the rank of the coefficient matrix and of the argument matrix, the reduced system can be written as equation (29). In that form, it can be seen that $x_{r+1}, x_{r+2}, \ldots, x_n$ can be given arbitrary values, say

$$x_{r+1} = \lambda_1, \quad x_{r+2} = \lambda_2, \quad \ldots, \quad x_n = \lambda_{n-r} . \tag{34}$$

With these values equation (29) reduces to

$$\begin{aligned}
x_1 &= \beta_1 - \alpha_{1,r+1}\lambda_1 - \alpha_{1,r+2}\lambda_2 - \cdots - \alpha_{1n}\lambda_{n-r} , \\
x_2 &= \beta_2 - \alpha_{2,r+1}\lambda_1 - \alpha_{2,r+2}\lambda_2 - \cdots - \alpha_{2n}\lambda_{n-r} , \\
&\vdots \\
x_r &= \beta_r - \alpha_{r,r+1}\lambda_1 - \alpha_{r,r+2}\lambda_2 - \cdots - \alpha_{rn}\lambda_{n-r} , \\
x_{r+1} &= \lambda_1 \\
x_{r+1} &= \lambda_2 \\
&\vdots \\
x_n &= \lambda_{n-r} .
\end{aligned} \tag{35}$$

or

$$\underset{\sim}{x} = \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_r \\ x_{r+1} \\ x \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \cdot \\ \cdot \\ \cdot \\ \beta_r \\ 0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{bmatrix} - \lambda_1 \begin{bmatrix} \alpha_{1,r+1} \\ \alpha_{2,r+1} \\ \cdot \\ \cdot \\ \cdot \\ \alpha_{r,r+1} \\ -1 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{bmatrix} - \lambda_2 \begin{bmatrix} \alpha_{1,r+2} \\ \alpha_{2,r+2} \\ \cdot \\ \cdot \\ \cdot \\ \alpha_{r,r+2} \\ 0 \\ -1 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{bmatrix} \tag{36}$$

$$+ \ldots - \lambda_{n-r} \begin{bmatrix} \alpha_{1n} \\ \alpha_{2n} \\ \cdot \\ \cdot \\ \cdot \\ \alpha_{rn} \\ 0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ -1 \end{bmatrix} .$$

tute equation (36) into equation (24), we get

$$A\left(\underset{\sim}{x}_p + \sum_{j=1}^{n-r} \lambda_j \underset{\sim}{x}_j\right) = A\underset{\sim}{x}_p + \sum_{j=1}^{n-r} \lambda_j A\underset{\sim}{x}_j \qquad (38)$$

$$= \underset{\sim}{b} + \underset{\sim}{0}$$

where

$$\underset{\sim}{x}_p = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \cdot \\ \cdot \\ \cdot \\ \beta_r \\ 0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{bmatrix} \qquad \underset{\sim}{x}_j = \begin{bmatrix} \alpha_{1,r+j} \\ \alpha_{2,r+j} \\ \cdot \\ \cdot \\ \cdot \\ \alpha_{r,r+j} \\ 0 \\ 0 \\ -1 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{bmatrix}$$

6. EXPANSION OF VECTORS IN A BASIS [2, 10]

A vector may be expanded in terms of a basis. This concept is very important in linear programming. We begin with the following definitions.

Spanning set: A set of vectors $\underset{\sim}{a}_1, \ldots, \underset{\sim}{a}_r$ from E^n is said to span or generate E^n if every vector in E^n can be written as a linear combination of $\underset{\sim}{a}_1, \ldots, \underset{\sim}{a}_r$.

Basis: A basis for E^n is a linearly independent subset of vectors from E^n which spans the entire space.

The representation of any vector $\underset{\sim}{b}$ in terms of a set of basis or a basis vectors is unique, that is, any vector in E^n can be written as a linear combination of a set of basis vectors in only one way. But, under certain conditions, an arbitrary vector, say b, from E^n can replace one of the vectors in a basis so that the new set of vectors is also basis. The technique of replacing one vector in a basis by another so that the new set is also a basis is fundamental to the simplex method for solving linear programming problems.

Consider a given set of basis vectors $\underset{\sim}{a}_1, \ldots, \underset{\sim}{a}_r$ for E^n and any other vector $\underset{\sim}{b} \neq \underset{\sim}{0}$ form E^n. Then b can be written as a linear combination of the $\underset{\sim}{a}_i$, as

$$\underset{\sim}{b} = \sum_{i=1}^{r} \alpha_i \underset{\sim}{a}_i \,. \tag{39}$$

where

$$a_i = \begin{bmatrix} a_{1i} \\ a_{2i} \\ \cdot \\ \cdot \\ \cdot \\ a_{mi} \end{bmatrix} \qquad b_i = \begin{bmatrix} b_1 \\ b_2 \\ \cdot \\ \cdot \\ \cdot \\ b_m \end{bmatrix}$$

If any vector $\underset{\sim}{a}_i$ for which $\alpha_i \neq 0$ is removed from the set $\underset{\sim}{a}_1, \ldots, \underset{\sim}{a}_r$ and $\underset{\sim}{b}$ is added to the set, the new collection of r vectors is also a basis for E^n.

By considering equation (39) as a system of simultaneous equations

$$\begin{aligned}
b_1 &= \alpha_1 a_{11} + \alpha_2 a_{12} + \ldots + \alpha_r a_{1r} , \\
b_2 &= \alpha_1 a_{21} + \alpha_2 a_{22} + \ldots + \alpha_r a_{2r} , \\
&\cdot \qquad \cdot \qquad\qquad\qquad \cdot \\
&\cdot \qquad \cdot \qquad\qquad\qquad \cdot \\
&\cdot \qquad \cdot \qquad\qquad\qquad \cdot \\
b_r &= \alpha_1 a_{r1} + \alpha_2 a_{r2} + \ldots + \alpha_r a_{rr} , \\
&\cdot \qquad \cdot \qquad\qquad\qquad \cdot \\
&\cdot \qquad \cdot \qquad\qquad\qquad \cdot \\
&\cdot \qquad \cdot \qquad\qquad\qquad \cdot \\
b_m &= \alpha_1 a_{m1} + \alpha_2 a_{m2} + \ldots + \alpha_r a_{mr} ,
\end{aligned} \tag{39}$$

$\{\alpha_1, \alpha_2, \ldots \alpha_r\}$ may be determined as discussed in section 5. Now consider the set of vectors with m components

$$\underset{\sim}{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \\ 0 \end{bmatrix}, \quad \underset{\sim}{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \\ 0 \end{bmatrix}, \ \ldots, \ \underset{\sim}{e}_m = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \\ 1 \end{bmatrix} \tag{40}$$

where $\underset{\sim}{e}_j$ is the vector with 1 in the jth position and zero elsewhere. Then we may express any vector in the original vector space consisting of $a_1, a_2, \ldots, a_n$ in terms of the set $\{e_j\}$ for

$$\underset{\sim}{a}_j = a_{1j}\, \underset{\sim}{e}_1 + a_{2j}\, \underset{\sim}{e}_2 + \ldots + a_{mj}\, \underset{\sim}{e}_m \quad \text{for all } j \tag{41}$$

and

$$\underset{\sim}{b} = b_1\, \underset{\sim}{e}_1 + b_2\, \underset{\sim}{e}_2 + \ldots + b_m\, \underset{\sim}{e}_m \,. \tag{42}$$

The set of vectors $\{e_j\}$ is called a set of unit vectors of the m-th

order and any vector with m components may be expressed or expanded in terms of this set. If we add to and subtract from equation (42) a constant multiple of one of the vectors of the set $\{a_j\}$, say, a_r, we obtain,

$$\underset{\sim}{b} = \sum_{i=1}^{m} b_i \underset{\sim}{e}_i - \alpha \underset{\sim}{a}_r + \alpha \underset{\sim}{a}_r \, . \tag{43}$$

Since

$$\underset{\sim}{a}_r = a_{1r}\underset{\sim}{e}_1 + a_{2r}\underset{\sim}{e}_2 + \ldots + a_{rr}\underset{\sim}{e}_r + \ldots + a_{mr}\underset{\sim}{e}_m \, , \tag{44}$$

we have

$$\underset{\sim}{b} = \sum_{i=1}^{m} (b_i - \alpha a_{ir})\underset{\sim}{e}_i + \alpha \underset{\sim}{a}_r \, . \tag{45}$$

Thus we obtain an expansion of the vector $\underset{\sim}{b}$ in terms of a new basis vector. This new basis includes $\underset{\sim}{a}_r$ and the original basis from which a vector is removed. The vector to be removed depending on the proper choice of α. If e_s is to be removed, we choose

$$\alpha = \frac{b_s}{a_{sr}} \, .$$

Then we have

$$\underset{\sim}{b} = \sum_{i=1}^{m} \left(b_i - \frac{b_s}{a_{sr}} a_{ir}\right)\underset{\sim}{e}_i + \frac{b_s}{a_{sr}} \underset{\sim}{a}_r \, . \tag{46}$$

Now each of the other vectors, not in the new basis, will have an expansion which is different from the original one. For instance, from equation (44), we can write

$$\underset{\sim}{e}_s = \frac{1}{a_{sr}}\left[\underset{\sim}{a}_r - a_{1r}\underset{\sim}{e}_1 - a_{2r}\underset{\sim}{e}_2 - \cdots - a_{s-1,r}\underset{\sim}{e}_{s-1}\right.$$

$$\left. - a_{s+1,r}\underset{\sim}{e}_{s+1} - \cdots - a_{my}\underset{\sim}{e}_m\right] \tag{47}$$

The expansion of any vector $\underset{\sim}{a}_j$ can be written from equations (41) and (47) as

$$\underset{\sim}{a}_j = \sum_{i=1}^{m} \left(a_{ij} - \frac{a_{sj}}{a_{sr}}\, a_{ir}\right) \underset{\sim}{e}_i + \frac{a_{sj}}{a_{sr}}\, \underset{\sim}{a}_r, \quad j \neq r$$

$$\underset{\sim}{a}_r = 0\underset{\sim}{e}_1 + 0\underset{\sim}{e}_2 + \cdots + 0\underset{\sim}{e}_m + \underset{\sim}{a}_r \,. \tag{48}$$

The procedure may be conveniently summarized in Table 6.1 and 6.2.

All the vectors under consideration are listed at the top of Table 6.1. Each vector consists of the components which are products of coefficients in the column below the vector and the corresponding unit vectors are listed in the first column.

Table 6.1

	$\underset{\sim}{a}_1$ ...	$\underset{\sim}{a}_j$ ...	$\underset{\sim}{a}_r$	$\underset{\sim}{e}_1$ ...	$\underset{\sim}{e}_s$ ...	$\underset{\sim}{e}_m$	$\underset{\sim}{b}$
$\underset{\sim}{e}_1$	a_{11}	a_{1j}	a_{1r}	1	0	0	b_1
$\underset{\sim}{e}_2$	a_{21}	a_{2j}	a_{2r}	0	0	0	b_2
$\underset{\sim}{e}_3$	a_{31}	a_{3j}	a_{3r}	0	0	0	b_3
.						.	
.						.	
.						.	
$\underset{\sim}{e}_s$	a_{s1}	a_{sj}	a_{sr}	0	1	0	b_s
.							.
.							.
.							.
$\underset{\sim}{e}_m$	a_{m1}	a_{mj}	a_{mr}	0 ...	0 ...	1	b_m

If we remove $\underset{\sim}{e}_s$ of the basis and introduce $\underset{\sim}{a}_r$ as a new basis and then we can determine the expansion coefficients in terms of the new basis by using equations (46), (47) and (48). This procedure is illustrated in Table 6.2. We shall now show that these entries in the table can be produced by operating on the rows of the matrix with elementary matrix operators. The entries in the row opposite $\underset{\sim}{a}_r$ in the first column of Table 6.2 may be obtained by dividing the row opposite $\underset{\sim}{e}_s$ in the first column of Table 6.1 by a_{sr}. Now note that all the elements under the column of $\underset{\sim}{a}_r$ in Table 6.2, except the one opposite $\underset{\sim}{a}_r$ in the first

column, are zero. For example, in order to produce a zero opposite $\underset{\sim}{e}_2$ in the first column, and under $\underset{\sim}{a}_r$ in the top row we divide the row opposite $\underset{\sim}{e}_s$ in Table 6.1 by a_{sr}, multiply it by a_{2r}, and subtract the result from the second row. This procedure produces the elements in the second row of Table 6.2. This simple numerical algorithm allows us to find expansions of vectors in terms of the new basis. Actually this is the same scheme as the Gauss-Jordon reduction mentioned in section 5.

Table 6.2

	$\underset{\sim}{a}_1$ ⋯	$\underset{\sim}{a}_j$ ⋯	$\underset{\sim}{a}_r$	$\underset{\sim}{e}_1$ ⋯	$\underset{\sim}{e}_s$ ⋯	$\underset{\sim}{e}_m$	$\underset{\sim}{b}$
$\underset{\sim}{e}_1$	$a_{11} - \frac{a_{s1}}{a_{sr}} a_{1r}$	$a_{1j} - \frac{a_{sj}}{a_{sr}} a_{1r}$	0	1	$-\frac{a_{1r}}{a_{sr}}$	0	$b_1 - \frac{b_s}{a_{sr}} a_{1r}$
.							
$\underset{\sim}{e}_2$	$a_{21} - \frac{a_{s1}}{a_{sr}} a_{2r}$	$a_{2j} - \frac{a_{sj}}{a_{sr}} a_{2r}$	0	0	$-\frac{a_{2r}}{a_{sr}}$	0	$b_2 - \frac{b_s}{a_{sr}} a_{2r}$
.	.	.					.
.	.	.					.
.	.	.					.
$\underset{\sim}{a}_r$	$\frac{a_{s1}}{a_{sr}}$	$\frac{a_{sj}}{a_{sr}}$	1	0	$\frac{1}{a_{sr}}$	0	$\frac{b_s}{a_{sr}}$
.	.						.
.	.						.
.	.						.
$\underset{\sim}{e}_m$	$a_{m1} - \frac{a_{s1}}{a_{sr}} a_{mr}$	$a_{mj} - \frac{a_{sj}}{a_{sr}} a_{mr}$	0	0	$-\frac{a_{mr}}{a_{sr}}$	1	$b_m - \frac{b_s}{a_{sr}} a_{mr}$

In order to illustrate the method, we may consider a simple example in which there are three basis vectors $\underset{\sim}{a}_1$, $\underset{\sim}{a}_2$, $\underset{\sim}{a}_3$, each with three components.

	$\underset{\sim}{a}_1$	$\underset{\sim}{a}_2$	$\underset{\sim}{a}_3$	$\underset{\sim}{e}_1$	$\underset{\sim}{e}_2$	$\underset{\sim}{e}_3$
$\underset{\sim}{e}_1$	3	6	3	1	0	0
$\underset{\sim}{e}_2$	1	2	3	0	1	0
$\underset{\sim}{e}_3$	3	2	0	0	0	1

We shall first remove $\underset{\sim}{e}_2$ replacing it with $\underset{\sim}{a}_3$ as shown below.

	$\underset{\sim}{a}_1$	$\underset{\sim}{a}_2$	$\underset{\sim}{a}_3$	$\underset{\sim}{e}_1$	$\underset{\sim}{e}_2$	$\underset{\sim}{e}_3$
$\underset{\sim}{e}_1$	2	4	0	1	-1	0
$\underset{\sim}{a}_3$	$\frac{1}{3}$	$\frac{2}{3}$	1	0	$\frac{1}{3}$	0
$\underset{\sim}{e}_3$	3	2	0	0	0	1

where we have operated on the rows of the first table to produce 1 at the intersection of the second row and third column and zeroes at all other positions of the third column. The second step is to replace $\underset{\sim}{e}_1$ with $\underset{\sim}{a}_2$. If we operate on rows to produce 1 at the intersection of the first row and second column, we obtain the following results.

	$\underset{\sim}{a}_1$	$\underset{\sim}{a}_2$	$\underset{\sim}{a}_3$	$\underset{\sim}{e}_1$	$\underset{\sim}{e}_2$	$\underset{\sim}{e}_3$
$\underset{\sim}{a}_2$	$\frac{1}{2}$	1	0	$\frac{1}{4}$	$-\frac{1}{4}$	0
$\underset{\sim}{a}_3$	0	0	1	$-\frac{1}{6}$	$\frac{1}{2}$	0
$\underset{\sim}{a}_3$	2	0	0	$-\frac{1}{2}$	$\frac{1}{2}$	1

Thus we can write the results as follow,

$$\underset{\sim}{a}_1 = \frac{1}{2}\,\underset{\sim}{a}_2 + 2\underset{\sim}{e}_3$$

$$\underset{\sim}{e}_1 = \frac{1}{4}\,\underset{\sim}{a}_2 - \frac{1}{6}\,\underset{\sim}{a}_3 - \frac{1}{2}\,\underset{\sim}{e}_3$$

$$\underset{\sim}{e}_2 = -\frac{1}{4}\,\underset{\sim}{a}_2 + \frac{1}{2}\,\underset{\sim}{a}_3 + \frac{1}{2}\,\underset{\sim}{e}_3$$

7. THE SIMPLEX METHOD [2, 9]

In Section 1, we discussed the geometrical interpretation of a linear programming problem in two dimensions. We observed that the set of feasible solutions to a linear programming problem described mathematically by equations (1) and (2) formed a convex region or convex set (if there a feasible solution). Furthermore, when an optimal solution exists, it is at one of the corner points or extreme points of the convex region of the feasible solution. It can be shown that in an n-dimensional convex region also, the optimum lies at one of the extreme points [2, 10]. Hence if we check also all the extreme points, then we shall be able to find the optimal solution. It may be proved [2, 6] that the extreme points of the convex set of the feasible solution are the basic feasible solution

to equation (8). Thus an optimal solution to a linear programming problem will be contained in the set of basic feasible solutions to equation (8).

The simplex method proceeds in systematic steps from an initial basic feasible solution to other basic solutions and finally in a finite number of steps to an optimal basic feasible solution, in such a way that the value of the objective function at each step (iteration) is better than it was at the preceeding step. Because the objective function z is improved at each step, the number of basic feasible solutions that must be examined before reaching an optimal solution is found, is usually much smaller than the total number of existing basic solutions. In general, the number of iteration (changes of basis) required to reach an optimal solution lies between m and 2m, where m is the number of constraints [10]. The simplex method also indicates whether there is an unbounded solution.

Now let us consider a linear programming problem with inequality constraints stated as maximizing

$$z = \underset{\sim}{c}^T \underset{\sim}{x} \tag{49}$$

$$A \underset{\sim}{x} \leq \underset{\sim}{b} \tag{50}$$

where

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdot & & & \\ \cdot & & & \\ \cdot & & & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

and

$$\underset{\sim}{x} = \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix} \qquad \underset{\sim}{b} = \begin{bmatrix} b_1 \\ b_2 \\ \cdot \\ \cdot \\ \cdot \\ b_m \end{bmatrix} \qquad \underset{\sim}{c} = \begin{bmatrix} c_1 \\ c_2 \\ \cdot \\ \cdot \\ \cdot \\ c_n \end{bmatrix}$$

By introducing slack variables into equation (9), as mentioned in section 4, it becomes

$$A_1 \underset{\sim}{x} = \underline{b} \tag{51}$$

subject to

$$z = \underset{\sim}{c}^T \underline{x} \tag{52}$$

where

$$A_1 = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} & 1 & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & a_{2n} & 0 & 1 & \cdots & 0 \\ \cdot & & & & & & & \cdot \\ \cdot & & & & & & & \cdot \\ \cdot & & & & & & & \cdot \\ a_{m1} & a_{m2} & \cdots & a_{mn} & 0 & 0 & \cdots & 1 \end{bmatrix}$$

and

$$\underset{\sim}{x} = \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_n \\ x_{n+1} \\ \cdot \\ \cdot \\ \cdot \\ x_{n+m} \end{bmatrix}, \quad \underset{\sim}{b} = \begin{bmatrix} b_1 \\ b_2 \\ \cdot \\ \cdot \\ \cdot \\ b_m \end{bmatrix}, \quad \underset{\sim}{c} = \begin{bmatrix} c_1 \\ c_2 \\ \cdot \\ \cdot \\ \cdot \\ c_n \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{bmatrix}$$

By defining

$$\underset{\sim}{P}_j = \begin{bmatrix} a_{1j} \\ a_{2j} \\ . \\ . \\ . \\ a_{ml} \end{bmatrix}, \quad j = 1,2, \ldots, N \text{ and } \underset{\sim}{P}_o = \begin{bmatrix} b_1 \\ b_2 \\ . \\ . \\ . \\ b_m \end{bmatrix}, \tag{53}$$

where

$$N = n + m$$

the linear program may be written as

$$\underset{\sim}{P}_0 = \sum_{j=1}^{N} x_j \underset{\sim}{P}_j$$

$$\max \underset{\sim}{c}^T \underset{\sim}{x}$$

$$\underset{\sim}{x} \geq 0 \tag{54}$$

We then ask what the coefficients x_j must be in the expansion of the vector $\underset{\sim}{P}_0$ in order that $\underset{\sim}{c}^T\underset{\sim}{x}$ be a maximum with $x \geq 0$. Since there are $N = m + n$ vectors in the expansion of a vector with m components, it is apparent that there can be many combinations of coefficients x_j, which

will satisfy the restraints. Note that only m vectors are required in a basis. Also note that $\underset{\sim}{P}_{n+1}$, $\underset{\sim}{P}_{n+2}$, ..., $\underset{\sim}{P}_N$ are m vectors which form a basis in a m-dimensional space and we may denote them as

$$\underset{\sim}{P}_{n+1} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{bmatrix}, \quad \underset{\sim}{P}_{n+2} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \\ 0 \end{bmatrix}, \quad \underset{\sim}{P}_N = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \\ 1 \end{bmatrix}. \tag{55}$$

It is now possible to find a basic feasible solution if we choose

$$\underset{\sim}{x}^T = \{0,0,\ldots 0, b_1, b_2, \ldots, b_m\} \tag{56}$$

where $\underset{\sim}{x}$ satisfies equality constraints

$$P_0 = \sum_{j=1}^{N} x_j \underset{\sim}{P}_j = \sum_{j=n+1} x_j \underset{\sim}{P}_j \tag{57}$$

because $x_j = 0$ for $j = 1, 2, \ldots, n$. In this case, the objective function

is

$$f = \underset{\sim}{c}^T \underset{\sim}{x} = \left[c_1, c_2, \ldots, c_n, 0, \ldots, 0\right] \begin{bmatrix} 0 \\ 0 \\ \vdots \\ b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} = 0 \qquad (58)$$

The reason for the objective function being zero is that only slack variables contribute to the objective function and they contribute "nothing" to it.

Recall that in section 6, we established a scheme whereby the basis can be changed by one vector at a time. If we change the basis by removing one vector from it and adding another one to it in a manner such that the profit function will always increase after a finite number of changes in the basis we should be able to arrive at the maximum value of z if it exists. The simplex method tells us which one of the vectors in the basis should be removed and which new vector should be added until we get the maximum value of z. In other words, it gives rise to the criterion

of optimality and provides us the direction of steepest decent. It may be developed as follows:

Suppose that after examining various combinations of $\underset{\sim}{P}_j$, we have arrived at the corresponding feasible solution $\underset{\sim}{x}$ with a set of basis vectors $\underset{\sim}{f}_1, \underset{\sim}{f}_2, \ldots, \underset{\sim}{f}_m$ where $\underset{\sim}{f}_i$ is an $\underset{\sim}{e}_j$, j=1, 2, ..., m or a $\underset{\sim}{P}_j$, j=1, 2, 3, ..., n. Let $\underset{\sim}{f}_{m+1}, \ldots, \underset{\sim}{f}_N$ be the remainder of the original N vectors. We employ $\underset{\sim}{f}_s$ to denote a vector belonging to the latter set. Since the vectors, $\underset{\sim}{f}_1, \underset{\sim}{f}_2, \ldots, \underset{\sim}{f}_m$, are linearly independent, they form a basis in an m-dimensional vector space. We can then express every vector as a linear combination of these basis vectors. Thus $\underset{\sim}{f}_s$ and $\underset{\sim}{P}_0$ can be expressed as

$$\underset{\sim}{f}_s = \sum_{j=1}^{m} \alpha_{js} \underset{\sim}{f}_j \tag{59}$$

and

$$\underset{\sim}{P}_0 = \sum_{j=1}^{m} \lambda_j \underset{\sim}{f}_j \, . \tag{60}$$

Define a quantity

$$z_s = \sum_{j=1}^{m} \lambda_{js} c_j \quad , \qquad s = m + 1, m + 2, \ldots m + n \tag{61}$$

and the corresponding values of z_s for the vectors $\underset{\sim}{f}_1 \underset{\sim}{f}_2, \ldots, \underset{\sim}{f}_m$ are

$$z_s = c_j \quad , \qquad j = s = 1, 2, \ldots, m. \tag{61-a}$$

and let the profit function be denoted by

$$z = \sum_{j=1}^{m} \lambda_j \, c_j \tag{62}$$

where the c_j in these formulas which belongs to c is the appropriate c_j associated with a particular $\underset{\sim}{f}_j$, j=1, 2, ..., n+m. If $\underset{\sim}{f}_j$ is a $\underset{\sim}{P}_i$, then c_j is c_i and on the other hand, if $\underset{\sim}{f}_j$ is an $\underset{\sim}{e}_k$ then c_j is zero.

Now suppose that one of $\underset{\sim}{f}_j$, j=1, 2, ..., m, is replaced in the basic basis by one of $\underset{\sim}{f}_s$, s = m+1, ..., N. Then

$$\underset{\sim}{P}_0 = \sum_{j=1}^{m} \lambda_j \, \underset{\sim}{f}_j - \theta \underset{\sim}{f}_s + \theta \underset{\sim}{f}_s$$

$$= \sum_{j=1}^{m} \lambda_j \, \underset{\sim}{f}_j - \theta \sum_{j=1}^{m} \alpha_{js} \, \underset{\sim}{f}_j + \theta \underset{\sim}{f}_s$$

$$= \sum_{j=1}^{m} (\lambda_j - \theta \alpha_{js}) \, \underset{\sim}{f}_j + \theta \underset{\sim}{f}_s \, .$$

$$= \sum_{i=1}^{m} (\lambda_i - \theta \alpha_{is}) \, \underset{\sim}{f}_i + \underset{\sim}{f}_s$$

If we chose

$$\theta = \frac{\lambda_j}{\alpha_{js}} \, ,$$

we may expand $\underset{\sim}{P}_0$ in terms of the new basis, that is, the basis with a particular $\underset{\sim}{f}_j$ removed and $\underset{\sim}{f}_s$ added to the set as shown below.

$$\underset{\sim}{P}_0 = \sum_{i=1}^{m} (\lambda_i - \frac{\lambda_j \alpha_{is}}{\alpha_{js}}) \underset{\sim}{f}_i + \frac{\lambda_j}{\alpha_{js}} \underset{\sim}{f}_s \ . \tag{63}$$

In the sum the term in which i = j is missing but the last term is substituted in its place. The vector λ^T, which is the solution of the problem now becomes

$$\underset{\sim}{\lambda}^T = (\lambda_i - \frac{\lambda_j \alpha_{is}}{\alpha_{js}}, \frac{\lambda_j}{\alpha_{js}}) \ . \tag{64}$$

The problem becomes that of choosing a proper in order to increase the objective function. In contrast to the profit function

$$z = \sum_{j=1}^{m} \lambda_j \ c_j \tag{65}$$

for the old set of $\underset{\sim}{f}_j$, the profit function for the new set of $\underset{\sim}{f}_j$ is

$$z_1 = \sum_{i=1}^{m} (\lambda_i - \frac{\lambda_j \alpha_{is}}{\alpha_{js}}) \ c_i + \frac{\lambda_j}{\alpha_{js}} \ c_s$$

$$= \sum_{i=1}^{m} \lambda_i \ c_i + \frac{\lambda_j}{\alpha_{js}} \ c_s - \sum_{i=1}^{m} \frac{\lambda_j \alpha_{is}}{\alpha_{js}} \ c_i$$

$$= \sum_{i=1}^{m} \lambda_i c_i + \frac{\lambda_j}{\alpha_{js}} c_s - \frac{\lambda_j}{\alpha_{js}} \sum_{i=1}^{m} \alpha_{is} c_i$$

$$= z + \frac{\lambda_j}{\alpha_{js}} (c_s - \sum_{i=1}^{m} \alpha_{is} c_i)$$

$$= z + \frac{\lambda_j}{\alpha_{js}} (c_s - z_s)$$

$$= z + \theta (c_s - z_s) \quad . \tag{66}$$

Thus the new value of the objective function increases by the quantity $\theta(c_s - z_s)$ if $\theta(c_s - z_s) > 0$.

Now suppose that θ has been chosen so that

$$\theta = \frac{\lambda_j}{\alpha_{js}} > 0 \quad , \tag{67}$$

implying that $\alpha_{js} > 0$ and also that

$$\lambda_i - \frac{\lambda_j}{\alpha_{js}} \alpha_{is} > 0, \text{ for all } i \neq j \ . \tag{68}$$

Then the solution of the problem given by equation (64)

$$\underset{\sim}{\lambda}^T = (\lambda_i - \frac{\lambda_j \alpha_{is}}{\alpha_{js}}, \frac{\lambda_j}{\alpha_{js}})$$

is a feasible solution. Note that $\lambda_j \geq 0$, therefore, the vector f_j to be removed in order to increase the objective function must be such that

$$(1) \quad \theta = \frac{\lambda_j}{\alpha_{js}} > 0 \tag{69}$$

or

$$\alpha_{js} > 0$$

$$(2) \quad \lambda_i - \frac{\lambda_j}{\alpha_{js}} \alpha_{is} > 0 \quad , i \neq j \tag{70}$$

or

$$\frac{\lambda_j}{\alpha_{js}} \text{ minum}$$

and the vector f_s to be added is chosen so that $(c_s - z_s)$ is the largest.

If all these conditions can be satisfied, a new feasible solution is obtained at an extreme point, and the process may then be repeated with successive improvements in the objective function. An iterative procedure may then be used until the maximum value of the objective function is obtained. Thus the procedure will terminate when values of all $(c_s - z_s)$ are either negative or zero.

In case all $\alpha_{is} < 0$, the procedure should also be terminate even though some of $(c_s - z_s)$ are still positive. This is the case of unbounded solution and no finite maximum exists. This case can be expressed as follows in the consideration of equations (66) and (63) repeated below

$$z_1 = \sum_{i=1}^{m} (\lambda_i - \frac{\lambda_j}{\alpha_{js}} \alpha_{is})\, C_i + \frac{\lambda_j}{\alpha_{js}} c_s$$

$$= \sum_{i=1}^{m} (\lambda_i - \theta\alpha_{is})\, c_i + \theta c_s$$

$$= z + \theta(c_s - z_s) ,$$

and

$$\underset{\sim}{P}_o = \sum_{i=1}^{m} (\lambda_i - \frac{\lambda_j}{\alpha_{js}} \alpha_{is})\, \underset{\sim}{f}_i + \frac{\lambda_j}{\alpha_{js}} \underset{\sim}{f}_s$$

$$= \sum_{i=1}^{m} (\lambda_i - \theta\alpha_{is})\, \underset{\sim}{f}_i + \theta \underset{\sim}{f}_s \quad .$$

Since α_{is} are negative and consequently all $(\lambda_i - \theta\alpha_{is})$ are non-negative and $\underset{\sim}{f}_i$ can be removed and $\underset{\sim}{f}_s$ can be introduced with a resulting feasible vector regardless of the value θ as long as it is positive for $\theta > 0$. Since some of $(c_s - z_s) > 0$ for in this case, it follows that z_1 can be made as large as possible by choosing θ to be as large as possible and therefore, the solution is unbounded and no finite maximum exists.

The method is best described by considering the so-called simplex tableau below.

	$\underset{\sim}{f}_1$	$\underset{\sim}{f}_2$	$\cdots$	$\underset{\sim}{f}_j$	$\cdots$	$\underset{\sim}{f}_m$	$\cdots$	$\underset{\sim}{f}_s$	$\cdots$	$\underset{\sim}{f}_N$	$\underset{\sim}{P}_0$	c_j
$\underset{\sim}{f}_1$	1	0	$\cdots$		$\cdots$	0	$\cdots$	α_{1s}	$\cdots$	α_{1N}	λ_1	c_1
$\underset{\sim}{f}_2$	0	1	$\cdots$		$\cdots$	0	$\cdots$	α_{2s}	$\cdots$	α_{2N}	λ_2	c_2
$\underset{\sim}{f}_j$	0	0	$\cdots$	1	$\cdots$	0	$\cdots$	α_{js}	$\cdots$	α_{jN}	λ_j	c_j
$\underset{\sim}{f}_m$	0	0	$\cdots$	0	$\cdots$	1	$\cdots$	α_{ms}	$\cdots$	α_{mN}	λ_m	c_m
c_s	c_1	c_2	$\cdots$	c_j	$\cdots$	c_m	$\cdots$	c_s	$\cdots$	c_N	—	—
z_s	z_1	z_2		z_j		z_m		z_s		z_N	z	—
$c_s - z_s$	$c_1 - z_1$	$c_2 - z_2$		$c_j - z_j$		$c_m - z_m$		$c_s - z_s$		$c_N - z_N$	—	—

Note that the values of $(c_s - z_s)$ corresponding to the vectors, $\underset{\sim}{f}_1$, $\underset{\sim}{f}_2$, ..., $\underset{\sim}{f}_m$ in the basis, are all zeros. Also note that, since the value of the objective function z is $\sum_{j=1}^{m} \lambda_j c_j$, it appears at the intersection of column $\underset{\sim}{P}_0$ and raw z_s.

Now we proceed to determine which vector should be removed and which should be introduced into the basis by means of this simplex tableau.

(1) Compute all $(c_s - z_s)$. Choose s so that $(c_s - z_s)$ is the largest positive number. This procedure fixes column s called the pivot column. The $\underset{\sim}{f}_s$ is to enter the new basis.

(2) Examine the column of $\underset{\sim}{f}_s$. Select those elements α_{is} which are greater than zero. Using these elements and the corresponding λ_i, compute

$\frac{\lambda_i}{\alpha_{is}}$. Choose the smallest $\frac{\lambda_i}{\alpha_{is}}$, say $\frac{\lambda_j}{\alpha_{js}}$. This procedure fixes the jth row called the pivot row.

(3) Use the method for replacing vectors in the basis in section (6). It can be implemented as follows:

(a) Divide the jth row by α_{js} to produce 1 in the column of $\underset{\sim}{f}_s$. at the jth row.

(b) Obtain zeros in all other rows of the table under $\underset{\sim}{f}_s$. For the ith row this step is done by multiplying the jth row by $\frac{\alpha_{is}}{\alpha_{js}}$ and subtracting the result element by element from the ith row.

(4) Compute $(c_s - z_s)$ for each column.

(5) If all $(c_s - z_s) \leq 0$, the vector $\underset{\sim}{\lambda}$ at this stage is optimal; if not, go back to step (1) and continue to interate.

(6) If one or more $(c_s - z_s) \geq 0$ and corresponding to these s, s = k for at least one of s with all $\alpha_{is} \leq 0$, then there exists an unbounded solution and thus the optimal solution does not exist.

Note that this procedure passes from one extreme point at a boundary of the admissible space to another until the objective function cannot be improved further by this method of computation.

8. GENERAL LINEAR PROGRAMMING PROBLEM

In section 2, we only considered the problem of linear programming with inequality constraints. In a more general case, the constraints may involve equalities. The general linear programming problem can be described as follows [10];

"Given a set of m linear inequalities or equalities in n variables, we wish to find non-negative values of these n variables which will satisfy the constraints and maximize or minimize some linear objective function of these n variables."

Mathematically, this statement means that we have m inequalities or equations in n variables (m can be greater than, less than, or equal to n) of the form:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \{\geq, =, \leq\} b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n \{\geq, =, \leq\} b_2 \\ \cdot \qquad\qquad \cdot \quad \cdot \\ \cdot \qquad\qquad \cdot \quad \cdot \\ \cdot \qquad\qquad \cdot \quad \cdot \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n \{\geq, =, \leq\} b_m \end{aligned} \tag{71}$$

This may be written in more compact form as

$$a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n \{\geq, =, \leq\} b_i \,, \tag{72}$$

$$i = 1, 2, \ldots, m \,,$$

or in matrix notation,

$$Ax \{\geq, =, \leq\} b \,, \tag{73}$$

where A is a matrix of coefficients of n variables in m constraints i.e.

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdot & & & \\ \cdot & & & \\ \cdot & & & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \tag{74}$$

$\underset{\sim}{x}$ is a column vector of n variables, i.e.

$$\underset{\sim}{x} = \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix} \tag{75}$$

and $\underset{\sim}{b}$ is also a column vector of m variables, i.e.

$$\underset{\sim}{b} = \begin{bmatrix} b_1 \\ b_2 \\ \cdot \\ \cdot \\ \cdot \\ b_m \end{bmatrix} \tag{76}$$

It should be noted that for each constraint one and only one of the signs $\leq$, $=$, $\geq$, holds but the sign may vary from one constraint to another. We seek values of the variables x_j satisfying (74) and

$$x_j \geq 0 \ , \ j = 1, 2, \ldots, n \ , \tag{77}$$

so that a linear objective function z defined below, is maximized or minimized

$$\begin{aligned} z &= c_1x_1 + c_2x_2 + \ldots + c_nx_n \\ &= \sum_{j=1}^{n} c_jx_j \end{aligned} \tag{78}$$

or in matrix notation

$$z = \underset{\sim}{c}^T \underset{\sim}{x}$$

where

$$\underset{\sim}{c}^T = \left(c_1 \ \ c_2, \ \ldots, \ c_n\right) .$$

We have thus formulated the general linear programming problem which, in short, can be written as follows:

to max or min $\quad z = \underset{\sim}{c}^T \underset{\sim}{x}$

subject to constraints (79)

$$\underset{\sim}{A}\underset{\sim}{x} \ \{\geq, =, \leq\} \ \underset{\sim}{b}, \qquad \underset{\sim}{x} \geq 0 \ .$$

Now we consider the transformation of the general linear programming problem. As mentioned in Section 4, the problem of inequality constraints with sign $\leq$ may be transformed to equality constraints by introducing slack variables. Specifically for the problem with inequality constraints

$$\sum_{j=1}^{n} a_{ij}x_j \leq b_i \ , \tag{80}$$

we may introduce slack variables

$$x_{n+i} \geq 0$$

where

$$x_{n+i} = b_i - \sum_{j=1}^{n} a_{ij}x_j$$

to change them to equality constraints given by

$$\sum_{j=1}^{n} a_{ij}x_j + x_{n+i} = b_i \ . \tag{81}$$

Now let us consider constraints with $\geq$ sign,

$$\sum_{j=1}^{n} a_{ij}x_j \geq b_i \ . \tag{82}$$

Here we introduce new variables x_{n+i} such that

$$x_{n+i} = \sum_{j=1}^{n} a_{ij}x_j - b_i \ .$$

These types of variables are called surplus variables. With introduction of the surplus variables inequalities are now changed to following equalities,

$$\sum_{j=1}^{n} a_{ij}x_j - x_{n+i} = b_i \ . \tag{83}$$

It is now obvious that we do not need any slack or surplus variables for equality constraints.

The original linear programming which has been stated as

$$\max z = \underset{\sim}{c}^T \underset{\sim}{x}$$

subject to,

$$\underset{\sim}{A}\underset{\sim}{x} \{\geq, =, \leq\} \underset{\sim}{b} \tag{84}$$

$$\underset{\sim}{x} \geq 0,$$

is transformed into

$$\max z = \underset{\sim}{c}^T \underset{\sim}{x}$$

subject to

$$\underset{\sim}{A}_1\underset{\sim}{x} = \underset{\sim}{b}$$

$$\underset{\sim}{x} \geq 0 \ , \tag{85}$$

Hence a column in A_1 in equation (85) corresponding to one of such surplus

variables is $-e_i$ because it has -1 as its coefficient as shown below,

$$A_1 = \begin{bmatrix} a_{11} & \cdots & a_{1n} & -1 & \cdots & 0 \\ \cdot & & & & \cdot\ \cdot & 0 \\ \cdot & & & & \cdot\ \cdot & \cdot \\ \cdot & & & & \cdot\ \cdot & \cdot \\ a_{m1} & & a_{mn} & 0 & & -1 \end{bmatrix}$$

In this case putting variables x_j equal to zero leads to

$$-I\underset{\sim}{s} = \underset{\sim}{b} \tag{89}$$

where s is a vector of surplus variables given by

$$\underset{\sim}{s} = \begin{bmatrix} x_{n+1} \\ x_{n+2} \\ \cdot \\ \cdot \\ \cdot \\ x_{n+m} \end{bmatrix}$$

we see that $\underset{\sim}{s}$ is not a feasible solution because all elements in it are negative. If we add, however, $\underset{\sim}{m}$ more variables called artificial variables x_{n+m+k}, $k = 1, 2, \ldots, m$, then we have column $\underset{\sim}{e}_i$ in A_2 which is defined below, and which corresponds to one of such artificial variables, because it has 1 as its coefficient.

$$\left[\begin{array}{ccc|ccc|ccc} a_{11} & \cdots & a_{1n} & -1 & \cdots & 0 & 1 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & & & & \vdots \\ \vdots & & \vdots & \vdots & \ddots & & & & \vdots \\ \vdots & & \vdots & \vdots & & \ddots & & & \vdots \\ a_{m1} & & a_{mn} & 0 & & -1 & 0 & \cdots & 1 \end{array}\right]$$

$$= \left[\begin{array}{c|c} A_1 & I \end{array}\right]$$

By letting both variables x_j, $j = 1, 2, \ldots, n$ and surplus variables x_{n+i}, $i = 1, 2, \ldots, m$, to be equal to zero, we have

$$I\underset{\sim}{r} = \underset{\sim}{b}$$

where $\underset{\sim}{r}$ is a vector of artificial variables x_{n+m+k}, $k = 1, 2, \ldots, m$. Note that r is a basic feasible solution to $A_2\underset{\sim}{x} = \underset{\sim}{b}$, but not a feasible solution to the original set of constraints $A_1\underset{\sim}{x} = \underset{\sim}{b}$. Any feasible solution to $A_2\underset{\sim}{x} = \underset{\sim}{b}$, which is also a feasible solution to the original set of constraints $A_1\underset{\sim}{x} = \underset{\sim}{b}$ must have $\underset{\sim}{r}=0$, that is, all artificial variables must vanish.

In the simplex method, $(z_s - c_s)$ determine which vector is to be inserted into the new basic in each interaction. Thus far we have said nothing about the prices to be associated with the artificial variables.

What we must do is to assign prices to the artificial variables, which are so unreasonable that the objective function can be improved as long as any artificial variable remains in the basic feasible solution. If z is to be maximized and if we assign an extremely large negative price to each artificial variable, we would expect that z can be improved so long as any artificial vector remains in the basis at a positive level. Similarly, if z is to be minimized, a very large positive price should be assigned to each artificial variable.

In case of equality constraints, we need to add neither slack variables nor surplus variables, but we need to add artificial variables to obtain an initial basic feasible solution.

Once we have an initial basic feasible solution, we may proceed to find an optimal one by simplex method.

9. APPLICATION OF LINEAR PROGRAMMING TO CHEMICAL ENGINEERING

The design and operation of chemical engineering processes gives rise to many optimization problems that can be solved by linear programming. In recent years problems of scheduling petroleum refinery operations [8, 14, 15, 16, 17], production and inventory control in a chemical process [15], chemical equilibrium [21], have been solved by linear programming. More specifically, linear programming has been used with excellent and very spectacular results on such problems as:

(1) Most profitable manufacturing problem.

(2) Best inventory strategies.

(3) Effect of changes in purchasing and selling price.

(4) Most profitable product mixture.

(5) Best location of plant.

(6) Best location of warehouses and distribution outlets.

(7) Lowest cost machine or manufacturing schedule.

Now, by means of several illustrative examples, we see how the simplex method can be used in solving optimization problems of chemical industry and chemical engineering which are amenable to the linear programming solutions.

(Example 1)

In chemical industrial operations, attempts are often made to maximize profits from available resources. Consider a chemical plant that has spare capacity in three batch-processing units in which we wish to make three possible products. Unit I is available for 22 hr./wk., Unit II for 14 hrs./wk., and Unit III for 14 hr/wk. Operating hours per ton of product for each product in each unit are given in the following table:

	Product A	Product B	Product C
Unit I	3	6	3
Unit II	1	2	3
Unit III	3	2	0
profit/ton, $	1000	4000	5000

We are required to find a production schedule that gives the maximum profit. This is an allocation-of-facility problem. Let

$$\underset{\sim}{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

be a production vector for products,

$$\underset{\sim}{c} = \begin{bmatrix} 1000 \\ 4000 \\ 5000 \end{bmatrix}$$

the profit vector; and

$$\underset{\sim}{b} = \begin{bmatrix} 22 \\ 14 \\ 14 \end{bmatrix}$$

the availability vector. Then the problem may be stated as follow.

Maximize the objective function,

$$f = 1000\, x_1 + 4000\, x_2 + 5000\, x_3$$

subject to constraints

$$3x_1 + 6x_2 + 3x_3 \leq 22,$$

$$x_1 \quad 2x_2 + 3x_3 \leq 14,$$

$$3x_1 + 2x_2 \leq 14,$$

$$x_j \geq 0, \qquad j = 1, 2, 3.$$

Let

$$c_1 = 1000, \quad c_2 = 4000, \quad c_3 = 5000,$$

and

$$P_1 = \begin{bmatrix} 3 \\ 1 \\ 3 \end{bmatrix}, \qquad P_2 = \begin{bmatrix} 6 \\ 2 \\ 2 \end{bmatrix}, \qquad P_3 = \begin{bmatrix} 3 \\ 3 \\ 0 \end{bmatrix},$$

Introducing slack variables x_4, x_5, and x_6, the constraints are transformed into equality constraints. Furthermore, letting

$$x_1 = \lambda_1, \quad x_2 = \lambda_2, \quad x_3 = \lambda_3, \quad x_4 = \lambda_4, \quad x_5 = \lambda_5,$$

$$x_6 = \lambda_6.$$

according to equation (60), we have

$$P_4 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \qquad P_5 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \qquad P_6 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

and

$$c_4 = 0, \qquad c_5 = 0, \qquad c_6 = 0 .$$

The initial simplex tableau may then be written as

(I)	$\underset{\sim}{P}_1$	$\underset{\sim}{P}_2$	$\underset{\sim}{P}_3$	$\underset{\sim}{P}_4$	$\underset{\sim}{P}_5$	$\underset{\sim}{P}_6$	$\underset{\sim}{P}_0$	c_j
$\underset{\sim}{P}_4$	3	6	3	1	0	0	22	0
$\underset{\sim}{P}_5$	1	2	(3)	0	1	0	14	0
$\underset{\sim}{P}_6$	3	2	0	0	0	1	14	0
c_s	1000	4000	5000	0	0	0	--	--
z_s	0	0	0	0	0	0	0	--
$c_s - z_s$	1000	4000	[5000]	0	0	0	--	--

Note that the feasible solution here is $x^T = [0, 0, 0, 22, 14, 14]$. Note also that the quantities of $(c_s - z_s)$ corresponding to the vectors in the basis, $\underset{\sim}{P}_4$, $\underset{\sim}{P}_5$, $\underset{\sim}{P}_6$, are all zero as mentioned in section 7 and those for the vectors not in the basis, we have $(c_s - z_s) = c_s$, since all $z_s = \sum_{j=1}^{3} \alpha_{js}\, c_j = 0$. The largest positive value of $(c_s - z_s)$ corresponds to $\underset{\sim}{P}_3$, that is, $s = 3$. Now we can proceed to construct the simplex tableau II according to the procedures described in section 7.

(1) Select the pivot column, in this case $\underset{\sim}{P}_3$, which is to enter the new basis.

(2) Examine the column of $\underset{\sim}{P}_3$, Compare

$$\frac{\lambda_4}{\alpha_{43}} \quad , \quad \frac{\lambda_5}{\alpha_{53}} \quad , \quad \frac{\lambda_6}{\alpha_{63}}$$

which are respectively

$$\frac{22}{3} \quad , \quad \frac{14}{3} \quad , \quad \frac{14}{0} \quad .$$

Choose the smallest one, in this case

$$\frac{\lambda_5}{\alpha_{53}} = \frac{14}{3}$$

This procedure fixes the pivot row to be P_5, which is to have the basis.

(3) Replace P_5 by the method described in section 6, which leads to tableau II.

(II)	P_1	P_2	P_3	P_4	P_5	P_6	P_0	c_j
P_4	2	(4)	0	1	-1	0	8	0
P_3	$\frac{1}{3}$	$\frac{2}{3}$	1	0	$\frac{1}{3}$	0	$\frac{14}{3}$	5000
P_6	3	2	0	0	0	1	14	0
c_s	1000	4000	5000	0	0	0	--	--
$z_s - z_s$	$\frac{5000}{3}$	$\frac{10000}{3}$	5000	0	$\frac{5000}{3}$	0	$\frac{70000}{3}$	--
$c_s - z_s$	$-\frac{2000}{3}$	$\boxed{\frac{2000}{3}}$	0	0	$-\frac{5000}{3}$	0	--	--

Now the feasible solution is $\underset{\sim}{x}^T = [\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5, \lambda_6] = [\ 0,\ 0,\ \frac{14}{3},\ 8,\ 0,\ 14\ ]$. The largest positive value of $(c_s - z_s)$, is 2000/3 corresponding to P_2, i.e. s = 2. Compute

$$\frac{\lambda_4}{\alpha_{42}} = \frac{8}{4} = 2, \quad \frac{\lambda_3}{\alpha_{52}} = \frac{\frac{14}{3}}{\frac{2}{3}} = 7, \quad \frac{\lambda_6}{\alpha_{62}} = \frac{14}{2} = 7.$$

Choose the smallest term which is $\frac{\lambda_4}{\alpha_{42}} = 2$, and corresponding to $\underset{\sim}{P}_4$. Replace $\underset{\sim}{P}_r$ by $\underset{\sim}{P}_2$. This procedure leads to Tableau III.

(III)	$\underset{\sim}{P}_1$	$\underset{\sim}{P}_2$	$\underset{\sim}{P}_3$	$\underset{\sim}{P}_4$	$\underset{\sim}{P}_5$	$\underset{\sim}{P}_6$	$\underset{\sim}{P}_0$	c_j
$\underset{\sim}{P}_2$	$\frac{1}{2}$	1	0	$\frac{1}{4}$	$-\frac{1}{4}$	0	2	4000
$\underset{\sim}{P}_3$	0	0	1	$-\frac{1}{6}$	$\frac{1}{2}$	0	$\frac{10}{3}$	5000
$\underset{\sim}{P}_6$	2	0	0	$-\frac{1}{2}$	$\frac{1}{2}$	1	10	0
c_s	1000	4000	5000	0	0	0	--	--
z_s	2000	4000	5000	$\frac{1000}{6}$	$\frac{3000}{2}$	0	$\frac{74000}{3}$	--
$c_s - z_s$	-1000	0	0	$-\frac{1000}{6}$	$-\frac{3000}{2}$	0	--	--

Since all $(c_s - z_s)$ are negative or zero, no further improvement of the objective function can be made. The optimal feasible solution is $\underset{\sim}{x}^T = [\ 0,\ 2,\ \frac{10}{3},\ 0,\ \frac{3}{2},\ 10]$ and the objective function,

$$f = 1{,}000x_1 + 4{,}000x_2 + 5{,}000x_3$$

$$= 0 + 4{,}000x_2 + 5{,}000x\frac{10}{3}$$

$$= \frac{74{,}000}{3}$$

$$= 24{,}333\,\frac{1}{3} \text{ dollars/week}$$

Therefore the optimal solution is

$$x_1 = 0, \quad x_2 = 2, \quad x_3 = \frac{10}{3}\ .$$

Note that the objective functions in Tableau I, II, and III are 0, $\frac{70{,}000}{3}$, $\frac{74{,}000}{3}$ respectively. The value increases step by step.

This problem is an original one worked out by the author.

(Example 2) [7]

In a chemical plant four raw materials A, B, C, D, are available and five processes are available to produce four products, E, F, G, H (Fig. 3). The raw materials are limited in supply and the processing cost of each process is fixed but varies from process to process. The selling price of the products is fixed. The data for the problem are summarized below:

Raw Materials	Maximum Available Supply (pounds per day)	Cost Per Pound (dollars)
A	400	1.50
B	300	2.00
C	100	4.50
D	250	2.50

The data on the processes are as follows:

Process	Pounds of Raw Material Used				Pounds (and type) of Product	Selling Price of Product (dollars per pound)
	A	B	C	D		
1	2	1	0	0	3 (E)	3.00
2	2	1	0	0	3 (F)	$2\frac{1}{3}$
3	3	1	0	2	6 (G)	3.75
4	2	7	3	3	15 (H)	5.00
5	2	7	3	3	15 (H)	5.00

Processes 4 and 5 differ only in processing cost. Process 4 has the maximum capacity of 75 lbs. of product H per day. If more than 75 lbs. of H are produced it must be produced in processes and processing costs are given below.

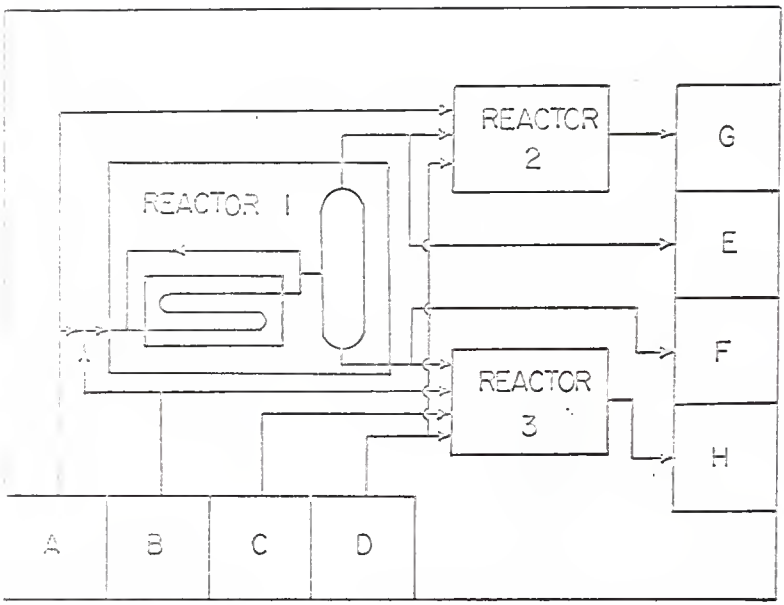

Fig. 3. Flow diagram of a chemical plant [7].

	Process	Processing Cost* (dollars per pound)
1	A + B → E	1.50 (A)
2	A + B → F	0.50 (A)
3	A + B → E	1.50 (A)
	E + A + D → G	1.00 (G)
4	A + B → F	0.50 (A)
	F + B + C + D → H	2.00 (H)
5	A + B → F	0.50 (A)
	F + B + C + D → H	2.20 (H)

*Based on the raw material or product indicated in the bracket

If we define x_j as the net profit from the jth process per day, the total profit per day can be written as

$$f = x_1 + x_2 + x_3 + x_4 + x_5 \tag{91}$$

The restrictions imposed on the processes because of the limitation of the raw material leads to the following formulation of inequality constrains.

Let A_i, B_i, C_i, and D_i be the amounts of the raw material A, B, C, and D used in process i. Then we can write,

$$A_1 + A_2 + A_3 + A_4 + A_5 \leq 400$$

$$B_1 + B_2 + B_3 + B_4 + B_5 \leq 300 \qquad (92)$$

$$C_1 + C_2 + C_3 + C_4 + C_5 \leq 100$$

$$D_1 + D_2 + D_3 + D_4 + D_5 \leq 250$$

$$H_4 \leq 75$$

where

$$C_1 = C_2 = C_3 = 0, \quad D_1 = D_2 = 0.$$

Let E_i, F_i, G_i, H_i, be the amount of product E, F, G, H, in the process i. The net profit of process 1, x_1, equal to the selling price of product E_1 subtracts the raw material cost and processing cost of the reactant A and B.

$$\begin{aligned} x_1 &= 3\ E_1 - 1.5\ A_1 - 2B_1 - 1.5A_1 \\ &= 3\ (\tfrac{3}{2} A_1) - 1.5\ A_1 - 2\ (\tfrac{1}{2} A_1) - 1.5\ A_1 \qquad (93) \\ &= 0.5\ A_1 \\ &= B_1 \end{aligned}$$

For the same reason,

$$x_2 = 2\tfrac{1}{3} F_2 - 0.5 A_2 - 2 B_2 - 1.5 A_2$$

$$= (2\tfrac{1}{3})\tfrac{3}{2} A_2 - 0.5 A_2 - 1.5 A_2 - 2(\tfrac{1}{2}) A_2 \qquad (94)$$

$$= 0.5 A_2$$

$$= B_2$$

$$x_3 = 3.75 G_3 - 1.0 G_3 - 2 B_2 - 2.5 D_3 - 1.5 A_3 - 1.5(\tfrac{2}{3}) A_3$$

$$= 3.75 G_3 - 1.5 A_3 - 2.0 B_3 - 2.5 D_3 - 1.5 A_3 - 1.0 G_3$$

$$= 3.75(2) A_3 - 1.5 A_3 - \tfrac{2}{3} A_3 - 2.5(\tfrac{2}{3} A_3) - 1.5 A_3 - 0.2 A_3$$

$$= \tfrac{2}{3} A_3 \qquad (95)$$

$$= 2 B_3$$

$$= D_3$$

$$x_4 = 5.0 H_4 - 1.5 A_4 - 2.0 B_4 - 4.5 C_4 - 2.5 D_4 - 0.5 A_4 - 2.0 H_4$$

$$= 5(\tfrac{15A_4}{2}) - 1.5 A_4 - 2.0(\tfrac{7A_4}{2}) - 4.5(\tfrac{3A_4}{2}) - 2.5(\tfrac{3A_4}{2}) - 0.5 A_4$$

$$- 2.0(\tfrac{15A_4}{2}) \qquad (96)$$

$$= 3 A_4$$

$$= \frac{6}{7} B_4$$

$$= 2 C_4$$

$$= D_4,$$

$$x_5 = 5.0 H_5 - 1.5 A_5 - 2.0 B_5 - 4.5 C_5 - 2.5 D_5 - 0.5 A_5 - 2.2 H_5$$

$$= 5(\frac{15A_4}{2}) - 1.5 A_4 - 20(\frac{7A_4}{2}) - 4.5(\frac{3A_4}{2}) - 2.5(\frac{3A_4}{2}) - 0.5 A_4$$

$$- 2.2(\frac{15A_4}{2}) \qquad (97)$$

$$= \frac{3}{2} A_5$$

$$= \frac{3}{7} B_5$$

$$= C_5$$

$$= D_5$$

Since process 4 has maximum capacity of 75 lbs. of H per day, we may write the constraint of x_4 as

$$x_4 = 3 A_4$$

$$= 3(\frac{2H_4}{15}) \qquad (98)$$

$$= \frac{2H_4}{5}$$

$$\leq \frac{2}{5}(75) = 30$$

By substituting equations (93) through (98) into equation (92), the inequality constraints become

$$\begin{aligned} 2x_1 + 2x_2 + \tfrac{3}{2}x_3 + \tfrac{1}{3}x_4 + \tfrac{2}{3}x_5 &\leq 400 \\ x_1 + x_2 + \tfrac{1}{2}x_3 + \tfrac{7}{6}x_4 + \tfrac{7}{3}x_5 &\leq 300 \\ \tfrac{1}{2}x_4 + x_5 &\leq 100 \\ x_3 + \tfrac{1}{2}x_4 + x_5 &\leq 250 \\ x_4 &\leq 30 , \end{aligned} \tag{99}$$

with

$$x_j \geq 0, \quad j = 1, 2, 3, 4, 5.$$

Introducing slack variables x_6, x_7, x_5, x_9, x_{10}, we may transform the preceeding inequality constraints into equality constraints. The problem may now be stated as follows

Maximize

$$f = x_1 + x_2 + x_3 + x_4 + x_5 \tag{100}$$

subject to constraints

$$\begin{aligned} 2x_1 + 2x_2 + \tfrac{3}{2}x_3 + \tfrac{1}{3}x_4 + \tfrac{2}{5}x_5 + x_6 &= 400 \\ x_1 + x_2 + \tfrac{1}{2}x_3 + \tfrac{7}{6}x_4 + \tfrac{7}{3}x_5 + x_7 &= 300 \\ \tfrac{1}{2}x_4 + x_5 + x_8 &= 100 \\ x_3 + \tfrac{1}{2}x_4 + x_5 + x_9 &= 250 \\ x_4 + x_{10} &= 30 \end{aligned} \tag{101}$$

with

$$x_j \geq 0 , \qquad j = 1, 2, \ldots, 10$$

Note that

$$c_1 = c_2 = c_3 = c_4 = c_5 = 1 ,$$

$$c_6 = c_7 = c_8 = c_9 = c_{10} = 0 ,$$

and

$$P_1 = \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad P_2 = \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad P_3 = \begin{bmatrix} \frac{3}{2} \\ \frac{1}{2} \\ 0 \\ 1 \\ 0 \end{bmatrix}, \quad P_4 = \begin{bmatrix} \frac{1}{3} \\ \frac{7}{6} \\ \frac{1}{2} \\ \frac{1}{2} \\ 1 \end{bmatrix}$$

$$P_5 = \begin{bmatrix} \frac{2}{3} \\ \frac{7}{3} \\ 1 \\ 1 \\ 0 \end{bmatrix}, \quad P_6 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad P_7 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad P_9 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix},$$

$$P_{10} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

The initial simplex tableau may then be written as,

(I)	$\underset{\sim}{P}_1$	$\underset{\sim}{P}_2$	$\underset{\sim}{P}_3$	$\underset{\sim}{P}_4$	$\underset{\sim}{P}_5$	$\underset{\sim}{P}_6$	$\underset{\sim}{P}_7$	$\underset{\sim}{P}_8$	$\underset{\sim}{P}_9$	$\underset{\sim}{P}_{10}$	$\underset{\sim}{P}_0$	c_j
$\underset{\sim}{P}_6$	(2)	2	$\frac{3}{2}$	$\frac{1}{3}$	$\frac{2}{3}$	1	0	0	0	0	400	0
$\underset{\sim}{P}_7$	1	1	$\frac{1}{2}$	$\frac{7}{6}$	$\frac{7}{3}$	0	1	0	0	0	300	0
$\underset{\sim}{P}_8$	0	0	0	$\frac{1}{2}$	1	0	0	1	0	0	100	0
$\underset{\sim}{P}_9$	0	0	1	$\frac{1}{2}$	1	0	0	0	1	0	250	0
$\underset{\sim}{P}_{10}$	0	0	0	1	0	0	0	0	0	1	30	0
c_s	1	1	1	1	1	0	0	0	0	0	--	-
z_s	0	0	0	0	0	0	0	0	0	0	0	-
$c_s - z_s$	[1]	1	1	1	1	0	0	0	0	0	--	-

The largest $(c_s - z_s)$ is 1 in column 1 and the smallest λ_i/α_{i1} is 200 row 6. So we replace $\underset{\sim}{P}_6$ by $\underset{\sim}{P}_1$ and obtain Tableau II given below

(II)	$\underset{\sim}{P}_1$	$\underset{\sim}{P}_2$	$\underset{\sim}{P}_3$	$\underset{\sim}{P}_4$	$\underset{\sim}{P}_5$	$\underset{\sim}{P}_6$	$\underset{\sim}{P}_7$	$\underset{\sim}{P}_8$	$\underset{\sim}{P}_9$	$\underset{\sim}{P}_{10}$	$\underset{\sim}{P}_0$	c_j
$\underset{\sim}{P}_1$	1	1	$\frac{3}{4}$	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{2}$	0	0	0	0	200	1
$\underset{\sim}{P}_7$	0	0	$-\frac{1}{4}$	1	$2 - \frac{1}{2}$		1	0	0	0	100	0
$\underset{\sim}{P}_8$	0	0	0	$\frac{1}{2}$	0	0	0	1	0	0	100	0
$\underset{\sim}{P}_9$	0	0	1	$\frac{1}{2}$	1	0	0	1	0	0	250	0
$\underset{\sim}{P}_{10}$	0	0	0	(1)	0	0	0	0	0	1	30	0
c_s	1	1	1	1	1	0	0	0	0	0	--	-
z_s	1	1	$\frac{3}{4}$	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{2}$	0	0	0	0	200	-
$c_s - z_s$	0	0	$\frac{1}{4}$	[$\frac{5}{6}$]	$\frac{2}{3} - \frac{1}{2}$		0	0	0	0	- 200	-

The largest $(c_s - z_s)$ is $\frac{5}{6}$ in column 4 and the smallest λ_i/α_{i4} in row 10. By replacing $\underset{\sim}{P}_{10}$ by $\underset{\sim}{P}_4$ we obtain Tableau III given below.

(III)	$\underset{\sim}{P}_1$	$\underset{\sim}{P}_2$	$\underset{\sim}{P}_3$	$\underset{\sim}{P}_4$	$\underset{\sim}{P}_5$	$\underset{\sim}{P}_6$	$\underset{\sim}{P}_7$	$\underset{\sim}{P}_8$	$\underset{\sim}{P}_9$	$\underset{\sim}{P}_{10}$	$\underset{\sim}{P}_0$	c_j
$\underset{\sim}{P}_1$	1	1	$\frac{3}{4}$	0	$\frac{1}{3}$	$\frac{1}{2}$	0	0	0	$-\frac{1}{6}$	195	1
$\underset{\sim}{P}_7$	0	0	$-\frac{1}{4}$	0	(2)	$-\frac{1}{2}$	1	0	0	0	70	0
$\underset{\sim}{P}_8$	0	0	0	0	1	0	0	1	0	$-\frac{1}{2}$	85	0
$\underset{\sim}{P}_9$	0	0	1	0	1	0	0	0	1	$-\frac{1}{2}$	235	0
$\underset{\sim}{P}_4$	0	0	0	1	0	0	0	0	0	1	30	1
c_s	1	1	1	1	1	0	0	0	0	0	-	-
z_s	1	1	$\frac{3}{4}$	1	$\frac{1}{3}$	$\frac{1}{2}$	0	0	0	$\frac{5}{6}$	225	-
$c_s - z_s$	0	0	$\frac{1}{4}$	0	$\boxed{\frac{2}{3}}$	$-\frac{1}{2}$	0	0	0	$-\frac{5}{6}$	- 225	-

The largest $(c_s - z_s)$ is $\frac{2}{3}$ in column 5 and the smallest λ_i/α_{i5} is 35 in rows 7. By replacing P_7 and P_5 we obtain Tableau IV given below.

(IV)	$\underset{\sim}{P}_1$	$\underset{\sim}{P}_2$	$\underset{\sim}{P}_3$	$\underset{\sim}{P}_4$	$\underset{\sim}{P}_5$	$\underset{\sim}{P}_6$	$\underset{\sim}{P}_7$	$\underset{\sim}{P}_8$	$\underset{\sim}{P}_9$	$\underset{\sim}{P}_{10}$	$\underset{\sim}{P}_0$	c_j
$\underset{\sim}{P}_1$	1	1	$\frac{19}{24}$	0	0	$\frac{7}{12}$	0	0	0	$-\frac{1}{6}$	$=\frac{550}{3}$	1
$\underset{\sim}{P}_5$	0	0	$-\frac{1}{8}$	0	1	$-\frac{1}{4}$	$\frac{1}{2}$	0	0	0	35	1
$\underset{\sim}{P}_8$	0	0	$\frac{1}{8}$	0	0	$\frac{1}{4}$	$-\frac{1}{2}$	1	0	$-\frac{1}{2}$	50	0
$\underset{\sim}{P}_9$	0	0	$\textcircled{\frac{9}{8}}$	0	0	$\frac{1}{4}$	$-\frac{1}{2}$	0	1	$-\frac{1}{2}$	200	0
$\underset{\sim}{P}_4$	0	0	0	1	0	0	0	0	0	1	30	1
c_s	1	1	1	1	1	0	0	0	0	0	-	-
z_s	0	0	$\frac{2}{3}$	1	1	$\frac{1}{3}$	$\frac{1}{2}$	0	0	$\frac{5}{6}$	$\frac{745}{3}$	-
$c_s - z_s$	0	0	$\boxed{\frac{1}{3}}$	0	0	$-\frac{1}{3}$	$-\frac{1}{2}$	0	0	$-\frac{5}{6}$	$-\frac{745}{3}$	-

The largest $(c_s - z_s)$ is $\frac{1}{3}$ in column 3, and the smallest λ_i/α_{i3} is $\frac{1600}{9}$ in row 9. By replacing $\underset{\sim}{P}_9$ by $\underset{\sim}{P}_3$ we obtain Tableau V given below

(V)	$\underset{\sim}{P}_1$	$\underset{\sim}{P}_2$	$\underset{\sim}{P}_3$	$\underset{\sim}{P}_4$	$\underset{\sim}{P}_5$	$\underset{\sim}{P}_6$	$\underset{\sim}{P}_7$	$\underset{\sim}{P}_8$	$\underset{\sim}{P}_9$	$\underset{\sim}{P}_{10}$	$\underset{\sim}{P}_0$	c_j
$\underset{\sim}{P}_1$	1	1	0	0	0	$\frac{11}{27}$	$\frac{19}{54}$	0	$-\frac{19}{27}$	$\frac{5}{27}$	$\frac{1150}{27}$	1
$\underset{\sim}{P}_5$	0	0	0	0	1	$-\frac{2}{9}$	$\frac{4}{9}$	0	$\frac{1}{9}$	$-\frac{1}{18}$	$\frac{515}{9}$	1
$\underset{\sim}{P}_8$	0	0	0	0	0	$\frac{2}{9}$	$-\frac{4}{9}$	1	$-\frac{1}{9}$	$-\frac{4}{9}$	$\frac{250}{9}$	0
$\underset{\sim}{P}_3$	0	0	1	0	0	$\frac{2}{9}$	$-\frac{4}{9}$	0	$\frac{8}{9}$	$-\frac{4}{9}$	$\frac{1600}{9}$	1
$\underset{\sim}{P}_4$	0	0	0	1	0	0	0	0	0	1	30	1
c_s	1	1	1	1	1	0	0	0	0	0	-	-
z_s	1	1	1	1	1	$-\frac{11}{27}$	$\frac{19}{54}$	0	$\frac{8}{27}$	$\frac{37}{54}$	$\frac{8305}{27}$	-
$c_s - z_s$	0	0	0	0	0	$-\frac{11}{27}$	$-\frac{19}{54}$	0	$-\frac{8}{27}$	$-\frac{37}{54}$	$-\frac{8305}{27}$	-

Since all $(c_s - z_s)$ are now negative or zero, no further improvement of the objective function can be made. The optimal solution is $x^T = [\ \frac{1150}{27}\ ,\ 0,\ \frac{1600}{9}\ ,\ 30,\ \frac{515}{9}\ ,\ 0,\ 0,\ \frac{250}{9}\ ,\ 0,\ 0\ ]$ and the corresponding value of the objective function is

$$f = x_1 + x_2 + x_3 + x_4 + x_5$$

$$= \frac{1150}{27} + \frac{1600}{9} + 30 + \frac{515}{9}$$

$$= \frac{8305}{27}$$

$$= 307\ \frac{16}{27}\ .$$

Therefore, the maximum attainable profit under the conditions of the problem is $307\ \frac{16}{27}$ dollars per day.

The original problem was solved by French and Acrivos [7] by using the simplex method. The author formulated the problem in a slightly different way that led to different simplex tableaus but eventually the same result was obtained.

(Example 3) [3, 16]

Four kinds of crude oil are available for purchase by an oil company: 100,000 barrels per week each of crudes 1, 2, and 3, and 200,000 barrels per week of crude 4. Let x_1 be the amount of crude 1 purchased (and processed), expressed in thousands of barrels per week.

Then the maximum availability of 100,000 bbls/wk of crude 1 implies that

$$x_1 \leq 100. \tag{103}$$

Similarly we can describe the constraints on crudes 2 and 3 by the inequalities:

$$x_2 \leq 100 \tag{104}$$

and

$$x_3 \leq 100 \tag{105}$$

where x_2, x_3 are defined in the same way as x_1.

Crude 4 requires slightly different handling because it can be processed two ways. Let x_4 be the amount of crude 4 processed to make heating oil, and let x_5 be the amount processed mainly to make lubricating oil. The constraint on crude 4 may be written as follow:

$$x_4 + x_5 \leq 200 \tag{106}$$

Four products - gasoline, heating oil, lubricating oil and jet fuel - are made from these crudes as shown by the block diagram of Fig. 4, Table 7.1 gives the amount of each product which can be sold. The bottom row of Table 7.1 shows the profit gained per 1000 barrels of crude processed. These numbers are obtained by adding up the market value of the products coming from 1000 barrels of the crude in question and then deducting the costs of production, sales, and the crude itself. The

TABLE 7.1 -- Crudes Used to Satisfy Product Needs [3]

	crude identity					Product on order, bbl/wk
	1	2	3	4		
	FUEL	FUEL	FUEL	FUEL	LUBE	
	Product Yield, Vol. %					
Gasoline	0.6	0.5	0.3	0.4	0.4	170,000
Heating oil	0.2	0.2	0.3	0.3	0.1	85,000
Lube oil	0	0	0	0	0.2	20,000
Jet fuel	0.1	0.2	0.3	0.2	0.2	85,000
Loss	0.1	0.1	0.1	0.1	0.1	
Total	1.0	1.0	1.0	1.0	1.0	
Crude available bbl/wk	100000	100000	10000	-200000-		-
Profit, $/M bbl crude	100	200	70	250		-

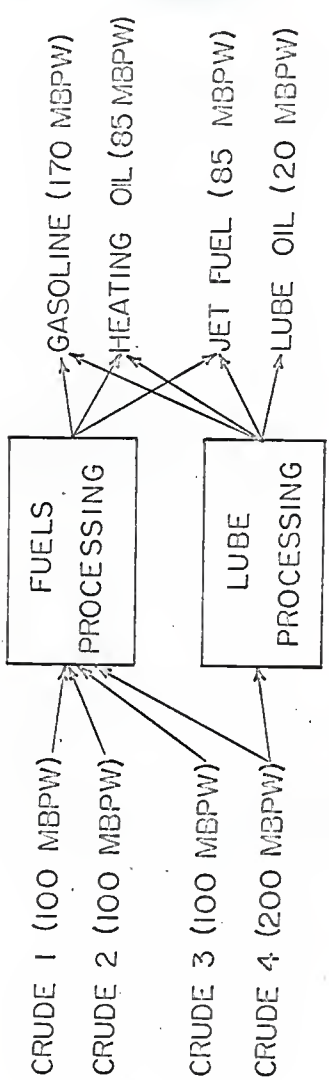

Fig. 4. Block diagrom of processing crude petroleum in o refinery [3].

yields are fixed by process technology and remain constant throughout the week. On the other hand, other data - availabilities, orders, and profits - are only estimates and may change between the receipt of the computer solution and the actual crude run.

The gasoline yields in the top row of Table 7.1 can be used to show that the weekly gasoline production (in thousands of barrels) in terms of the crude consumption is

$$0.6x_1 + 0.5x_2 + 0.3x_3 + 0.4x_4 + 0.4x_5.$$

It will be assumed that we are permitted to make less product than is ordered, but no more. Thus we may write the following constraint on gasoline production.

$$0.6x_1 + 0.5x_2 + 0.3x_3 + 0.4x_4 + 0.4x_5 \leq 170 \tag{107}$$

Similarly we obtain production constraints on heating oil,

$$0.2x_1 + 0.2x_2 + 0.3x_3 + 0.3x_4 + 0.1x_5 \leq 85, \tag{108}$$

lube oil,

$$0.2x_5 \leq 20 \tag{109}$$

and jet fuel,

$$0.1x_1 + 0.2x_2 + 0.3x_3 + 0.2x_4 + 0.2x_5 \leq 85. \tag{110}$$

The bottom line of Table 7.1 enables us to calculate the profit, f, in dollars per week as

$$f = 100x_1 + 200x_2 + 70x_3 + 150x_4 + 250x_5 .$$

The optimal running plan would be the set of feasible x's which maximize this profit.

By introducing slack variable x_6, x_7, ..., x_{13} into equations (103), (104), ..., (110) respectively. All the inequality constraints may be transformed into equality constraints and this problem may be stated as follow:

$$f = 100x_1 + 200x_2 + 70x_3 + 150x_4 + 250x_5 \tag{111}$$

subject to constraints

$$\begin{aligned}
x_1 + x_6 &= 100 \\
x_2 + x_7 &= 100 \\
x_3 + x_8 &= 100 \\
x_4 + x_5 + x_9 &= 200 \\
0.6x_1 + 0.5x_2 + 0.3x_3 + 0.4x_4 + 0.4x_5 + x_{10} &= 170 \\
0.2x_1 + 0.2x_2 + 0.3x_3 + 0.3x_4 + 0.1x_5 + x_{11} &= 85 \\
0.2x_5 + x_{12} &= 20 \\
0.1x_1 + 0.2x_2 + 0.3x_3 + 0.2x_4 + 0.2x_5 + x_{13} &= 85
\end{aligned} \tag{112}$$

$x_i \geq 0$, for $i = 1, 2, \ldots, 13$.

The initial simplex tableau may then be written as

(I)	P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8	P_9	P_{10}	P_{11}	P_{12}	P_{13}	P_0	c_j
P_6	1	0	0	0	0	1	0	0	0	0	0	0	0	100	0
P_7	0	1	0	0	0	0	1	0	0	0	0	0	0	100	0
P_8	0	0	1	0	0	0	0	1	0	0	0	0	0	100	0
P_9	0	0	0	1	1	0	0	0	1	0	0	0	0	200	0
P_{10}	0.6	0.5	0.3	0.4	0.5	0	0	0	0	0	0	0	0	170	0
P_{11}	0.2	0.2	0.3	0.3	0.1	0	0	0	0	0	1	0	0	85	0
P_{12}	0	0	0	0	0.2	0	0	0	0	0	0	1	0	20	0
P_{13}	0.1	0.2	0.3	0.4	0.2	0	0	0	0	0	0	0	1	85	0
c_s	100	200	70	150	250	0	0	0	0	0	0	0	0	-	-
z_s	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-
$c_s - z_s$	-100	-200	-70	-150	-250	0	0	0	0	0	0	0	0	-	-

Starting with the initial simplex tableau, we may solve the problem by the simplex method. The problem was originally solved by Beighter and Wilde. The author solved it by using a 1410 computer. The optimal solution is found to be

$x_1 = 37.5$, $x_2 = 100$, $x_3 = 58, 33$ $x_4 = 100$ $x_5 = 100$.

and the maximum profit is

$$f = 100x_1 + 200x_2 + 70x_3 + 150x_4 + 250x_5$$

$$= 100 \times 37.5 + 200 \times 100 + 70 \times 58.33 + 150 \times 100 + 250 \times 100$$

$$= 67,833.33 \text{ dollars per week.}$$

10. CONCLUDING REMARK

Linear programming is one of the most powerful and most frequently used optimization methods. Problems with 25,000 variables can now be solved on the computer by using the simplex method [12, 13]. However the linear programming algorithms are only valid for linear problems, and the majority of process design problems are more accurately represented by nonlinear design relations (1).

A compromise must be reached between the use of a powerful linear optimization method on an approximate linear model of a chemical process, and the use of the less efficient nonlinear programming methods on a more accurate model of the process. In many cases, we are forced to abandon the linear programming, since the important features the engineer wishes to incorporate into his optimum design are inherently nonlinear. To linearize the equations describing most chemical processes is to do irreparable damage to reality (13).

ACKNOWLEDGMENTS

The author wishes to express his gratitude to his major advisor, Dr. Liang-tseng Fan, for his excellent help and constant encouragement.

The interest and suggestions of the members of the committee, Dr. William H. Honstead, Dr. Larry E. Erickson, and Dr. E. Stanley Lee are deeply appreciated.

The author also wishes to thank Dr. Frank A. Tillman for his advice concerning the computation of the simplex method of linear programming.

The financial aids provided by the Kansas State University Engineering Experiment Station and project 2409 supported by the Office of Saline Water are also acknowledged.

REFERENCES

PART I

1. Asbjornsen, O. A., "Incomplete Mixing Simulated by Fluid-Flow Networks," a paper presented at A.I.Ch.E. - I.Ch.E. Joint Meeting, 39, London, June, (1965).

2. Danckwerts, P. V. "Continuous Flow System-- Distribution of Residence Time," Chem. Eng. Sci., 2, 1 (1953).

3. Danckwerts, P. V., "The Effect of Incomplete Mixing on Homogeneous Reactions," Chem. Eng. Sci. 8, 93 (1958).

4. Denbigh, K. G., "Instaneous and Overall Reaction Yields," Chem. Eng. Sci., 14, 25 (1961).

5. Douglas, J. M., "The Effect of Mixing on Reactor Design," Chem. Eng. Progr. symposium ser., 60, 1 (1964).

6. Kattan, A., "A Mixing Model of Continuous Flow Chemical Reactor," Ph.D. Thesis, Case Inst. of Techno., Cleveland, Ohio (1967).

7. Kattan, A. and R. J. Adler, "A Mixing Model for Continuous Chemical Reactor," Research Report No. 02-10-67, Chem. Eng. Sci. Group, Case Inst. of Techno. Cleveland, Ohio (1967).

8. Kattan, A., and R. J. Adler, "A Stochastic Mixing Model for Homogeneous, Turbulent, Tubular Reactors," Research Report No. 07-05-66, Chem. Eng. Sci. Group, Case Inst. of Techno., Cleveland, Ohio.

9. Kramers, I. H., "Physical Factors in Chemical Reaction Engineering," Chem. Eng. Sci. 8, 45 (1958).

10. Kranback, F. J., Shinnar, R., and S. Katz, "Stochastic Mixing Models for Chemical Reactors," I&EC Fundamentals, 6, 276 (1967).

11. Makoto, H., Kikuo, H., Wataru, E., and Shinji, N., "Micromixing in a Continuous Flow Reactor (Coalescence and Redispersion Models), Memoirs of the Faculty of Eng., 14, Part 4, Kyoto University, Kyoto, Japan (1962).

12. Ng, D. Y., and D. W. T. Rippin, "The Effect of Incomplete Mixing on Conversion in Homogeneous Reactions," Proc. 3rd Europ. Symposium of Chem. Eng., Amsterdam, Netherlands (1964).

13. Rietema, K., "Segregation in Liquid-Liquid Dispersions and Its Effect on Chemical Reactions," Adv. in Chem. Eng., Edited by T. B. Drew, Vol. 5, 237, Academic Press (1964).

14. Rippin, D. W. T., "The Recycle Reactor as a Model of Incomplete Mixing," Manuscript from I&EC Research Results Services, 66-348-J.

15. Rippin, D. W. T., "Segregation in a Two-Environment Model of a Partially Mixed Chemical Reactor," Chem. Eng. Sci., 22, 247 (1967).

16. Fukuzumi, S., "Macromixing and Dynamics of a Stirred Tank Reactor," M. S. thesis, Tokyo Inst. of Techno., Tokyo, Japan (1966).

17. Weinstein, H., "The Effect of Mixing on Conversion in Continuous Chemical Reactors," Ph.D. thesis, Case Inst. of Techno., Cleveland, Ohio (1963).

18. Weinstein, H. and R. J. Adler, "Micromixing Effects in Continuous Chemical Reactors," Chem. Eng. Sci., 22, 65 (1967).

19. Worrel, G. R., and L. E. Eagleton, "An Experimental Study of Mixing and Segregation in a Stirred Tank Reactor," Can. J. Chem. Eng., 39, 254 (1964).

20. Zwietering, Tr. N., "An Experimental Study of Mixing and Segregation," Chem. Eng. Sci., 11, 1 (1959).

21. Cha, L., "Analysis of Flow Systems and Their Applications to Reactor Design," Ph.D. thesis, Kansas State University, Manhattan, Kansas (1965).

PART II

1. Acrivos, A., "Linear Programming--How Does It Work?" Chem. Eng., 63, 8 Aug. (1956).

2. Amunsdon, N. R., "Mathematical Methods in Chemical Engineering," Prentice-Hall, 1966.

3. Beighter, C. S. and D. J. Wilde, "Sensitivity Analysis Gives Better Insight Into Linear Programming," Hydrocarbon Processing, 44, 2, (1965).

4. Charles, R. N., "How to Formulate a Linear Programming Model for Refinery Simulation" The Oil and Gas Journal, 57, 7 (1959).

5. Dantiz, G. B., "Thoughts on Linear Programming and Automation" Manag. Sci., 32 2 (1957).

6. Dantiz, G. B., "Linear Programming and Extensions", Princeton University Press, 1963.

7. French, E. J., and A. Acrivos, "The Application of Linear Programming to Design Problems," Chem. Eng. Sci., 5, 93, (1956).

8. Garvin, W. W., etc., "Applications of Linear Programming in the Oil Industry," Manag. Sci., Oct. (1957).

9. Gass, S. I., "Linear Programming," McGraw-Hill, 1964.

10. Hadley, G., "Linear Programming" Addison-Wesley, 1962.

11. Llewellyn, R. W., "Linear Programming," Hlt, Rinehart and Winston, (1964).

12. Naylor, T. H. and E. T. Byrne, "Linear Programming," Wadsworth, 1963.

13. Rudd, D. F. and C. C. Watson, "Strategy in Process Engineering", to be published, John Wiley, 1966.

14. Schwan, H. T. and J. J. Wilkinson, "Linear Programming--What Can It Do," Chem. Eng., 63, 8, Aug. (1956).

15. Stillson, P., "Operation Research--Application in the Chemical Industries." Ind. Eng. Chem. 48, 3 (1956).

16. Symonds, G. H., "Linear Programming: The Solution of Refinery Problems," Esso Standard Oil Company (1955).

17. Symonds, G. H., "Linear Programming Solves Gasoline Refining and Blending Problems," Ind. Eng. Chem. 48, 3 (1956).

18. Van Canwenberghe, A. R., "An Optimization Application of Linear Programming," British Chem. Eng. 2, 12 (1964).

19. Wilde, D. J., "Production Planning of Large Systems", Chem. Eng. Prog., 59, 1 (1963).

20. Wilde, D. J., "Optimization Methods," Adv. in Chem. Eng. 3, Academic Press (1962).

"EFFECT OF MIXING ON REACTOR PERFORMANCE" AND "APPLICATION OF LINEAR PROGRAMMING TO CHEMICAL ENGINEERING"

by

Billy I Tsai

B. S., National Taiwan University, 1963

AN ABSTRACT OF A MASTER'S REPORT

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Chemical Engineering

KANSAS STATE UNIVERSITY

Manhattan, Kansas

1968

ABSTRACT

In part one, an extensive review of literature concerning the effect of mixing on the reactor performance is presented. Two aspects of mixing, micromixing and macromixing are fully discussed. A detailed example of the effect of mixing under extreme conditions of the conversion is given for an isothermal reactor. Various models of incomplete mixing are also reviewed.

In part two, development of the simplex method of linear programming is presented. Typical examples are given to illustrate the application of linear programming in chemical industries and processes.