The data given in this report are of importance when considering the utilization of carbohydrate residues. In the industrial hydrolysis of starches and agricultural residues to sugars, the production of large quantities of 5-hydroxymethyl-2-furaldehyde and levulinic acid results in low yields of the sugars. However, the potential importance of these two materials as chemical intermediates makes them valuable products themselves. This is especially true in relation to the chemical utilization of wood.

The kinetics of the decomposition of glucose in aqueous acidic media and its conversion to HMF and levulinic acid is known for over a century. It has been rather well established by several investigators (1, 2, 3) that the mechanism of the reaction consists essentially of a series of consecutive reactions which follow this order:

\[ \text{Glucose} \rightarrow \text{CH}_3\text{COH}_2\text{CH}_2\text{COOH} \]

(Levulinic acid)

\[ + \]

\[ \text{CHOOH} \]

(Formic acid)
They suggest that the reaction proceeds through several intermediate compounds and that one mole each of HMF levulinic acid, and formic acid is produced per mole of initial glucose.

The overall kinetics of this series of reactions has not been developed in the past, but several of the individual steps have been investigated. The disappearance of glucose in dilute solutions of mineral acid has been followed (4, 1, 5) at temperatures below 200°C, and the disappearance of HMF in dilute mineral acids has also been investigated (6, 1, 7). It has further been shown (1, 8) that levulinic acid displays no marked tendency to decompose at these reaction conditions.

Most of the previous work in this field, however, was concerned primarily with the determination of yields of HMF and levulinic acid produced from various sugars and the reaction conditions necessary to obtain these yields. The variations of the yields with time, as are necessary in a kinetics study, were in general not included. In all cases, the yields were less than the theoretical, and it was concluded that the reaction mechanism was complicated.

**Experimental and Analytical Techniques**

This investigation was carried out in two main studies. In the first, the disappearance of glucose and the simultaneous formation of HMF and levulinic acid was measured. Measurements were also taken on the formation of the total organic acids and the insoluble solid materials. In the second, the disappearance of HMF and the simultaneous formation of levulinic acid was determined. The range of variables for these studies included temperatures of 140°C to 250°C, catalyst acid concentrations of 0.025 N to 0.8 N sulfuric acid, and several concentrations of initial glucose and HMF.

The experimental data were obtained in batch reactors which were heated in a constant temperature bath for various time intervals. The reactors were glass ampoules made from 5-millimeter (Pyrex) glass tubing which had a capacity of 1.0 milliter of solution. Mechanical strength tests showed that the ampoules could withstand internal pressures of approximately 1,300 pounds per square inch without failing, and heat transfer tests showed that the ampoules' contents reached the bath temperature in 25 seconds. An appropriate correction factor was applied to all the results to compensate for this short warm-up period.

Ampoules were charged with precise amounts of solution from a hypodermic syringe pipe, were chilled to reduce the vapor pressure, and then a vacuum applied to remove oxygen and reduce the pressure in the tube. While the vacuum was still being applied, the tubes were sealed off and the ampoule was completed. This loading technique showed variations in the volume charged of less than 0.2 percent and was rapid enough to allow preparation of 200 ampoules per day.

Ampoules were reacted in a thermostatically controlled bath of hydrogentated cottonseed oil for specific time intervals and were then quenched in a second bath. Transfer of ampoules from one bath to the other was performed behind safety shields with the aid of rotating brackets with control rods. The heating bath was completely enclosed to contain any explosions resulting from the occasional failure of an ampoule. The temperature of the bath was controlled within 0.1°C at the lower temperatures and 0.2°C at the higher temperatures.

The contents of the ampoules were obtained for analysis by breaking open both ends and transferring the reacted solutions to vials. The analysis for glucose was that described by Shaffer and Somogyi (9) and depends on the reducing power of glucose. The reacted solutions were analyzed directly for residual glucose since there were no reaction products present which would interfere. The reducing power of HMF was found to be 5 percent of that of glucose, and it was present in too small quantities to influence the measurements.

Due to the presence of materials which would interfere in subsequent analyses, the HMF was first chromatographically separated on paper from the reaction mixture. The solvent system employed was the organic layer of pentanol. A standardization of the procedure with known quantities of HMF showed that its recovery from the paper was quantitative at 92 percent with a standard deviation for the experiment.
ment points of 1.59-percent. The indicated 8 percent loss is probably due to its volatility since tests showed that less than 1 percent of the applied HMF was located outside the band eluted for analysis. The solutions of separated HMF were then analyzed for optical absorbance at 284 mµ, with a Beckman model DU spectrophotometer to determine the concentration.

An investigation of the optical properties of the known and suspected reaction products and their behavior with the solvent system mentioned, showed that they either would not separate with HMF or they would not absorb significantly at 2M mµ. Also, a study of the optical spectrum of HMF separated from reaction mixtures did not indicate the presence of contaminants. Due to a tendency of the dilute HMF solutions to degrade to 5-hydroxymethyl-2-furoic acid at room temperatures and when exposed to light, it was necessary to store them in a dark refrigerator until they could be analyzed.

The levulinic acid analysis was also preceded by a paper chromatographic separation. In this separation, the solvent system consisted of ethanol (95 percent), aqueous ammonia (29 percent), and water in a volume ratio of 100 :5 :5, respectively. Due to its volatility with this solvent, the levulinic acid was chromatographed as its ammonium salt which separated into a narrow, distinct band. After elution of this band, it was in turn converted to the sodium salt and the ammonium ions removed by evaporating to dryness at room temperature since the ammonium ions would interfere with the subsequent analysis. The method employed to measure the sodium levulinate concentration of the separated solution was suggested by Ploetz and Bartels (10) and depends on the acetyl group, which undergoes the iodoform reaction in the presence of alkaline iodine solutions. The stoichiometry of this reaction has not been satisfactorily developed, and hence a calibration was made with pure levulinic acid under conditions which were stringently observed in subsequent work. This calibration indicated that 7.18 equivalents of iodine are consumed per equivalent of levulinic acid. The standard deviation of the experimental points in this calibration was 1.74 percent.

The analysis for the total organic acids present consisted of titrating the reacted samples with a standardized base and subtracting the acidity of the catalyst acid, and the amount of solids formed was found by filtering, drying and weighing the solids from a reacted ampoule.

Glucose Disappearance

It was found that the glucose disappearance follows a first-order mechanism:

\[
\ln C_G^e = \frac{C_G}{C_G^0} = e^{-k_1 t}
\]  

Experimental values of \(\ln C_G^e\) plotted against time resulted in straight lines from which the values of \(k_1\) were determined. A sample plot is shown in Figure 1. The rate constants were correlated as prescribed by the Arrhenius equation from which an activation energy of 32,690 gram-calories per gram-molecule was obtained.

To incorporate the proportional dependency of the rate constants on the catalyst acid concentration, an activity term was included in the coefficient of the Arrhenius equation. The standard state for the activity coefficient, \(a_G\), was arbitrarily defined as unity at 180°C, a catalyst concentration of 0.8 N., and an initial glucose concentration of 0.556 gram-molecule per liter. The values of the activity coefficients at other conditions were then calculated using the experimental values of \(k_1\).

Over the range of temperatures and acid concentrations investigated, \(a_G\) was only a function of catalyst concentration and showed only minor, random variations with temperature. This relationship is shown in Figure 2. It should be noted that the activity coefficient, although it resembles that for sulfuric acid, is defined for the solution as a whole, hence, the subscript \(G\).

Most of the experimental work was carried out at an initial glucose concentration of 0.556 gram-molecule per liter with a few additional runs at 0.278 and 1.116 gram-molecule per liter to determine the effect of varying the initial glucose concentration. At all concentrations a first-order disappearance occurred; that is, the plots of \(\ln C_G^e\) versus time were all linear. However, as seen in Figure 1,
the slopes of these lines and, consequently, the values of $k_1$ were not the same for various concentrations. It is held, nevertheless, that the reaction is a true first-order and the discrepancy is attributed to the manner in which the solutions were made up. This apparent discrepancy is discussed in a later section. It was consequently necessary to include the initial glucose concentration as a parameter in the activity coefficient plot.

The corrected Arrhenius equation has this form:

$$k_1 = (9.27 \times 10^{15}) (a C_A) e^{-\frac{E_A}{RT}}$$

(2).

It may also be noted that the glucose disappearance curves have intercepts at time equal to zero that decrease as the concentration increases. This is probably due to the reversion of the glucose to polysaccharides which has been shown to be more pronounced at higher concentrations. This does not affect the slopes and, consequently, the rate constants, however.

**HMF Formation from Glucose**

The formation of HMF from glucose shows a growth and decay with time as typified by Figure 3. The data are reported as moles present at any time per mole of initial glucose, $C_i$, and represent the fraction of the theoretical yield. The time function is a dimensionless time, $k_1 t$, obtained by multiplying the glucose disappearance rate constant by the isothermal reaction time. This term was used because it allowed all the data to be shown on one graph. It is related to the glucose disappearance half-life, $t_{1/2}$, thus:

$$k_1 t = 0.69$$

In Figure 4 are plotted the maximum yields of HMF where it is shown that the yields are increased with increased temperature and decreased initial glucose concentration, but are independent of catalyst concentration except at higher temperatures where they are increased with decreasing amounts of arid. A possible explanation of this effect is given later. The times at which these maxima occur may be calculated using equation (4) below and will be seen to shift to longer reaction times as the yields increase and vary from one-third to one and one-third glucose half-lives. These yields may also be converted to moles of HMF formed per mole of glucose reacted with the aid of equation (1). On this basis, the maximum HMF yields vary from 0.18 to 0.27 moles/mole.

If one neglects the presence of intermediates and assumes first-order reactions exclusively, the following simplified mechanism can be written:

$$\text{Glucose} \rightarrow \text{HMF} \rightarrow \text{arid}$$

Assuming constant volume and the boundary condition that the HMF concentration, $C_H$, equals zero at time equals zero, the dependence of $C_H$ with time is:

$$C_H^* = \frac{C_H}{C_G} = \frac{1}{a - 1} \left[ e^{-(k_1 t)} - e^{-(a k_1 t)} \right]$$

(3)

Table 2.—MAXIMUM YIELDS OF TOTAL ORGANIC ACIDS FROM GLUCOSE ($C_G^* = (\text{moles/mole})$)

<table>
<thead>
<tr>
<th>Temperature (°C)</th>
<th>0.8</th>
<th>0.4</th>
<th>0.2</th>
<th>0.1</th>
<th>0.05</th>
</tr>
</thead>
<tbody>
<tr>
<td>180</td>
<td>1.18</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>190</td>
<td>1.13</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>200</td>
<td>1.18</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>220</td>
<td>1.10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>240</td>
<td>0.85</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$C_G = 0.556$ gram-molecules per liter except as noted.

$C_G = 0.278$ gram-molecules per liter.

$C_G = 1.112$ gram-molecules per liter.
Fig. 7.—Shown is the effect of sulfuric acid concentration on the yield of total organic acids produced from glucose \( (C_G = 0.556 \text{ gram-molecules per liter, } 200^\circ \text{C.}) \).

Fig. 8.—Solids formation from glucose is shown.

At catalyst concentrations less than 0.1 N., however, this relationship no longer holds.

The correlating equation was further modified by the inclusion of a constant coefficient to yield this final form:

\[
(C_L^*) = \left( \frac{0.51}{A} \right) \left( \frac{1}{1 - e^{-\frac{(k_1 t)}{a}}} \right)
\]

(6).

This equation is shown plotted in Figure 3 where it can be seen that it accurately describes the experimental data.

**Levulinic Acid Formation From Glucose**

The formation of levulinic acid from glucose displays a plateau value, as typified in Figure 6, due to the stability of levulinic acid to these reaction conditions. The data are reported as moles of levulinic acid per mole of initial glucose, \( C_L^* \), and since the yield reaches its plateau values, \( (C_L^*)_p \), at approximately six glucose half-lives, the plateau value is also the moles of levulinic acid per mole of glucose reacted. Mechanisms proposed by previous investigators suggest that we mole is formed per mole of glucose. From the size of the yields it can be inferred that there are complicating side reactions. The time in this case is also \( (k_1 t) \). It was found found \( (C_L^*)_p \) increases with decreased temperature and decreased initial glucose concentration and is independent of catalyst concentration except at low concentrations where it decreases with decreased amounts of catalyst. Values for \( (C_L^*)_p \) are given in Table 1.

The simple two-step mechanism considered earlier can be extended to include levulinic acid where only first-order reactions are considered:

Glucose \( \xrightarrow{k_1} \) HMF \( \xrightarrow{k_2} \) levulinic acid.

It was found in this investigation that the levulinic acid displayed no tendency to decompose and, therefore, no degradation is included above. Assuming constant volumes and the boundary condition that \( C_H^* \) and \( C_L^* \) are zero when time is zero, the dependence of \( C_L^* \) with time is:

\[
C_L^* = C_L^0 - \left( \frac{a}{a - 1} \right) \left[ 1 - e^{-\frac{(k_1 t)}{a}} \right] - 1 \left( 1 - e^{-\frac{(k_1 t)}{a}} \right)
\]

(7).

As with the HMF data, it was found that the reactions involved were too complicated to lend themselves to an exact discussion of their mechanisms, and consequently an empirical approach was used. Equation (7) was modified by substituting the correlating function \( A \) discussed earlier for \( a \) and multiplying the right side by the plateau value of \( C_L^* \), \( (C_L^*)_p \). This latter modification was necessary since the values of \( C_L^* \) did not approach unity at large values of \( k_1 t \) as demanded by equation (7), but reached lower values which depended on the reaction conditions.
With the values of \((C_L)_p\) from Table 1, equation (8) is shown plotted with the experimental data in Figure 6. A good fit is obtained not only at the plateau, as is expected, but also between time zero and the plateau. In extending the values of \((C_L)_p\) to other reaction conditions, it should be noted that ln \((C_L)_p\) plotted versus \((1/T)\) using initial glucose concentration as the parameter results in a linear plot which is independent of catalyst concentration for values of \(C_A\) greater than 0.1 N.

Organic Acid and Solids Formation From Glucose

The data on the formation of organic acids from glucose are reported as moles of organic acids per mole of initial glucose, \(C_H^0\). To do this, it was assumed that all the acids present were single dissociation acids since it was believed that formic and levulinic acids were the main acid materials present. The yield curves as a function of time display a maximum with only a slight tendency to decay as typified by Figure 7. This is probably predominantly due to the disappearance of the formic acid which is formed. Earlier mechanisms suggested that one mole each of formic and levulinic acids were formed per mole of initial glucose. It was found that the yields were considerably below two, as predicted by this assumption, and that they were also greater than twice as large as the yield of levulinic acid. The values of the maxima are listed in Table 2 where it will be noted that the maximum yields are increased with decreased temperature, decreased initial glucose concentration, and increased catalyst concentration. Since the maxima do not occur until six glucose half-lives, these values are also essentially moles formed per mole of glucose reacted.

Approximately 25 percent of the initial glucose is converted to solid materials during the course of the reaction. The formation curves when plotted versus \((k_1t)\) are a function only of the initial glucose concentration and are independent of temperature and catalyst concentration over the range investigated (180°C.–220°C.) and 0.2 N.—0.4 N. sulfuric acid). These data are shown plotted in Figure 8.

5-Hydroxymethyl-2-Furaldehyde Disappearance and Simultaneous Appearance of Levulinic Acid

The HMF disappearance followed a first-order mechanism; that is, plots of \(\ln (C_H^0/C_H)\) versus time yielded straight lines, where \(C_H^0\) is the initial concentration of HMF. The HMF disappearance rate constants, \(k_2\), obtained from the slopes of these lines were correlated with temperature by means of the Arrhenius equation and an activation energy of 23.110 gram-calories per gram-molecule was determined.

To incorporate the dependency of the rate constant on catalyst concentration, an activity coefficient, \(a_H\), was included in the coefficient of the Arrhenius equation which was arbitrarily defined as unity at 180°C. and \(C_A = 0.8\) N. From the experimental value of \(k_2\) at these conditions, the numerical form of the equation was determined:

\[
k_2 = (2.40 \times 10^{11}) (a_H C_A) e^{(-23.110/RT)}
\]  

(9).

The values of \(a_H\) were then calculated at the other conditions from the experimental values of \(k_2\). Over the range of variables investigated, \(a_H\) was a function only of catalyst concentration and showed only minor, random variations with temperature. This relation between \(a_H\) and \(C_A\) is shown in Figure 2. As with the function, \(a_O\), it should be noted the \(a_H\) is defined for the entire solution, hence the subscript \(H\).

Most of the experimental work in this study was done at an initial HMF concentration of 0.0805 gram = molecules per liter, since this approximately the concentration of the HMF in the solutions in the glucose disappearance study. When the concentration was varied from 0.061 to 0.139 gram = molecules per liter, only minor variations were noted in the values of \(k_2\).

Table 3.—PLATEAU VALUES OF LEVULINIC ACID FROM 5-HYDROXYMETHYL-2-FURALDEHYDE (moles/mole) 

<table>
<thead>
<tr>
<th>Sulfuric acid concentration, (C_A), gram-equivalent per liter</th>
<th>0.02</th>
<th>0.01</th>
<th>0.05</th>
<th>0.025</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temperature °C.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>160</td>
<td>0.85</td>
<td>0.86</td>
<td>0.84</td>
<td>0.72</td>
</tr>
<tr>
<td>180</td>
<td>0.83</td>
<td>0.83</td>
<td>0.73</td>
<td>0.61</td>
</tr>
<tr>
<td>200</td>
<td>0.66</td>
<td>0.66</td>
<td>0.52</td>
<td></td>
</tr>
<tr>
<td>220</td>
<td>0.52</td>
<td>0.52</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

These data were not comprehensive enough, however, to determine if the disappearance were truly first-order or pseudo first-order and the solutions were too dilute to be expected to show the variation with initial concentration which was found in the glucose disappearance study.

The formation of levulinic acid from HMF displays a plateau value as a function of time. These plateau yields, shown in Table 3, display the same general trends with temperature and catalyst concentration indicated for levulinic acid previously. The mole fraction yields, however, are approximately twice as large in this study, but still less than the theoretical. The effect of varying the initial HMF concentration was not fully determined nor was a correlation justified in this study.

Discussion

The experimental data are correlated with empirical equations based on a simple first-order consecutive reaction mechanism. Because this mechanism itself does not adequately describe the true process, it is necessary to incorporate correlating functions to account for the deviations from this assumption. When the curves predicted by the equations are compared to the experimental data, however, it must be concluded that the trends are accurately described.

Although it would normally be desirable to use equations based on the true mechanism, the empirical approach was necessary in this study because the reactions are very complex and there was not sufficient information available to adequately determine the true reaction phenomena. Moreover, even if the actual mechanism were known, the equations necessary to describe a process of this complexity would, in themselves, be so complicated as to seriously hinder their application in any subsequent design work. For example, the differential equations which describe even the relatively sim-
ple mechanism given below are nonlinear.

The effect of the initial glucose concentration on the rate of glucose disappearance is attributed to the manner in which the solutions were made up. The solutions were prepared by weighing into a volumetric flask the desired amount of glucose, adding the desired amount of 1.00 N. sulfuric acid, and diluting with water. This procedure neglects the change in density of the solutions due to the glucose concentration and, hence, the acid to water ratio is higher in the solutions of higher glucose concentration, although the normality remains the same. The change in rate may be explained by a change in activity of the acid as the acid to water ratio varies.

This is further verified by the work of Kirby (11) in which the glucose disappearance rate constants did not vary over an eight-fold change in initial glucose concentration. His solutions were made up using different weights of glucose and sulfuric acid of the same normality. This resulted in solutions of varying final acid normalities, but in a constant ratio of acid to water. The solutions were made up in this work as described, however, because it was felt at the time that maintaining the normality constant was fundamentally more desirable.

It was also shown that the glucose disappearance was approximately proportional to the catalyst acid concentration. Figure 2 shows that at concentrations greater than 0.1 N. they were almost directly proportional, but at lower values they were not. In addition, yields of HMF and levulinic acid also show a similar dependence. This may be explained in part by considering each of the rate constants involved as composed of three parts: an acid catalyzed constant, an uncatalyzed constant, and a base catalyzed constant. At the higher acid concentrations, any variation in concentration may not cause a large enough change in the pH or in the ratio of the basic constant to the over-all constant to be noticeable, but at the lower concentrations, the changes in pH are large enough to cause this ratio to become significantly larger. Since different rate constants are involved in the production of each compound, however, it would not be expected that the variations in their yields would show the same dependency.

It was also found that the yields of HMF increase as the initial glucose concentration is decreased. This would indicate that HMF disappears by a higher order reaction. It appears most likely that it reacts with one of its precursors in this reaction rather than with itself or with one of its reaction products, since the HMF disappearance study indicated a first-order decay for HMF itself. A possible mechanism is:

\[
\text{Glucose} \xrightarrow{k_1} \xrightarrow{k_2} \text{HMF} \xrightarrow{k_3} \text{levulinic acid}
\]

where \( I \) is an intermediate; \( k_1, k_2, \) and \( k_3 \) are first-order constants; and \( k_4 \) is a higher order rate constant. A mechanism of this type would also be responsible for the decrease in levulinic acid yields as the initial glucose concentration is increased and for the general lowering of the yields from the theoretical value.

The solids have been shown as the probable result of the interaction between \( I \) and HMF since the data on solids formation indicate that its formation occurs by a higher order mechanism. Most of the data collected on solids formation are shown in Figure 8. The sulfuric acid concentrations are all 0.2 N. or 0.4 N. and show no influence on the amount of solids produced. In several experiments at lower acid concentrations, however, the yields of solids were greater. It seems probable, then, that the formation of the solids is responsible for the decrease in levulinic acid yields in a mechanism similar to that shown above, and that this side reaction becomes significantly more important at decreasing catalyst concentrations.

The results of the second study indicate that the HMF disappearance also appears to follow a first-order mechanism, even when the reaction was carried to 85 percent of completion. The yields of levulinic acid in this study were almost twice as high when obtained from glucose, again pointing out the presence of side reactions in the chain between glucose and HMF. The yields of levulinic acid in this study were less than the theoretical, however, indicating that side reactions or equilibria also exist between HMF and levulinic acid. These plateau values also showed the trend to become more strongly influenced by the catalyst concentration at low concentrations and higher temperatures.

The above mechanism is obviously an oversimplification of the true reaction path, but does help to explain some of the observed phenomena. There undoubtedly are other intermediates in the over-all reaction as well as other side reactions. Although none of these intermediates were isolated and identified, their presence was demonstrated. The chromatographic separations employed in the HMF and levulinic acid analyses both showed the presence of several unaccounted for compounds in appreciable concentrations.

**NOTATION:**

\[ A = \text{correlating function (dimensionless)} \]
\[ C = \text{concentration (gram = molecules per liter)} \]
\[ C^* = \text{dimensionless concentration (C/C_a)} \]
\[ C_a = \text{catalyst concentration (gram = equivalent per liter)} \]
\[ HMF = 5\text{-hydroxymethyl-2-furaldehyde} \]
\[ k_1 = \text{first-order glucose disappearance rate constant (minutes}^{-1}) \]
\[ k_2 = \text{first-order 5-hydroxymethyl-2-furaldehyde disappearance rate constant (minutes}^{-1}) \]
\[ t = \text{time (minutes)} \]
\[ a = \text{solution activity coefficient} \]

**SUBSCRIPTS:**

\[ A = \text{catalyst acid (sulfuric acid)} \]
\[ G = \text{glucose} \]
\[ G_p = \text{glucose present initially} \]
\[ H = 5\text{-hydroxymethyl-2-furaldehyde} \]
\[ H_p = 5\text{-hydroxymethyl-2-furaldehyde present initially} \]
\[ L = \text{levulinic acid} \]
\[ m = \text{maximum} \]
\[ OA = \text{organic acids} \]
\[ p = \text{plateau} \]
\[ S = \text{insoluble solid materials} \]

**Literature Cited:**