The UMAP Journal

The Journal of Undergraduate Mathematics and its Applications

Breaking the Missouri License Code

Joseph A. Gallian

COMAP

Spring 1992

1

Practitioner's Commentary: Optimization in Chemical Kinetics

Arthur T. Benjamin
Mathematics Dept.
Gordon J. Hogenson
Harvey Mudd College
Claremont, CA 91711

Introduction

In their paper, "A Mathematical Solution to a Chemical Engineering Problem," James Meehan, Edward Ferroni, and Laura Budrik solve a chemical engineering problem using numerical methods [Meehan et al. 1992]. We present alternative approaches to the same problem.

We are given an initial chemical solution of X-X molecules, and we choose how many Y molecules to add to the solution. A Y molecule may permanently attach to either end of an X-X molecule, resulting in an X-X-Y molecule (X-X-Y and Y-X-X molecules are identical in structure). Alternatively, a Y molecule may permanently attach to the "open" X on an X-X-Y molecule, forming a Y-X-X-Y molecule. Our goal is to add the number of Y molecules which maximizes the resulting number of X-X-Y molecules produced.

One Molecule at a Time

Our first approach is to imagine that we start with some large initial supply of X-X molecules, say $x_0 = 1,000$, and that we add Y molecules to the solution, at a rate of one molecule per unit of time. Hence we may replace y in our discussion by its equal, the elapsed time t.

After adding a number t of Y molecules to the solution, let x_t, z_t, and w_t denote, respectively, the number of X-X, X-X-Y, and Y-X-X-Y molecules present. (Note that if $t \leq 2x_0$, then there will be no free Y molecules in the solution.) Also, observe that $x_t + z_t + w_t = x_0$.

We now drop in another Y molecule. What is the probability that it attaches itself to an X-X molecule?

If we assume that a given X-X molecule is just as likely as a given X-X-Y molecule to attach itself to the Y molecule, then this probability would be $x_t/(x_t + z_t)$.

Perhaps, however, it's more natural to assume, since the X-X molecule has two open ends, that a given X-X molecule should be twice as likely as a given X-X-Y molecule to capture the newly-arrived Y molecule. Under this assumption, our probability becomes $2x_t/(2x_t + z_t)$.

Our model for the general problem discussed by Meehan et al. is to assume that a Y molecule is r times as likely to be attracted to a particular X-X molecule than to a particular X-X-Y molecule, where $r = k_1/k_2$ is the ratio of the reaction rate constants. Hence, the probability that our $(t+1)^{\text{st}}$ Y molecule attaches itself to an X-X molecule is $p_t = rx_t/(rx_t + z_t)$.

A simulation approach to this problem would go as follows. Initially, set $x_0 = 1{,}000, z_0 = 0$. Now, for $t \geq 0$, we proceed iteratively. Generate a random real number u between 0 and 1. If $u \leq p_t$, then our $(t+1)^{\text{st}}$ Y molecule attaches itself to an X-X molecule; and we set $x_{t+1} = x_t - 1$ and $z_{t+1} = z_t + 1$. Otherwise, it attaches itself to an X-X-Y molecule; and we set $x_{t+1} = x_t$ and $z_{t+1} = z_t - 1$. We observe the time(s) t when z_t is at its maximum value. Then we repeat this process several times, each time observing the t-values when the number of X-X-Y molecules is maximized. Each run suggests how many Y molecules we should add to our solution. If enough runs are performed, we might get an idea about the optimal number of Y molecules to add, relative to the number of initial X molecules.

One Mole at a Time

There are certain obvious drawbacks to the above approach. One is that it is not clear what to do when, on the same run, we have more than one t-value that maximizes z_t. Averaging these t-values may not be appropriate. A more serious drawback, however, is that it is impractical to imagine depositing one Y molecule at a time. More realistically, we should imagine starting with some large quantity (say 1,000 *moles*) of substance X-X, and trying to determine how many moles of Y should be added. With the assistance of a theorem from statistics, we can accurately predict how many moles of Y should be added.

As before, we continue to assume that when we drop in our $(t+1)^{\text{st}}$ mole of Y, then (as before) all molecules (an Avogadro number, 6.02×10^{23}, of them) instantly and independently attach themselves to either an X-X molecule (with probability p_t) or an X-X-Y molecule (with probability $1 - p_t$). By the law of large numbers from statistics, we can expect the fraction p_t of our mole to react with X-X molecules and the remainder to react with X-X-Y molecules. Hence, if at time t we have x_t moles of X-X and z_t moles of X-X-Y, then, after adding a mole of Y, we can expect to have $x_{t+1} = x_t - p_t$ and $z_{t+1} = z_t + p_t - (1 - p_t) = z_t + 2p_t - 1$, where $p_t = rx_t/(rx_t + z_t)$. Thus, we have the following difference equations

$$x_{t+1} = x_t - \frac{rx_t}{rx_t + z_t}, \qquad z_{t+1} = z_t + \frac{rx_t - z_t}{rx_t + z_t}.$$

Using these equations, we can determine when z_t is maximized, by using a simple computer program. We begin with $x_0 = 1{,}000$ moles of X-X and

$z_0 = 0$ moles of X-X-Y. Then we iterate, incrementing t from 0 to 2,000 (when all the molecules have become Y-X-X-Y molecules) and note the time t when z_t is maximized. The maximizing t-value is the number of moles of Y to use for every 1,000 moles of X-X. The results are presented in **Table 1**, where $y^* = t^*$ is the number of moles of Y (per 1,000 moles of X-X) that maximizes the amount of X-X-Y, for our difference-equation model.

Table 1.

Results of the "one mole at a time" method.

r	y^*	x^*	z^*
0.002	19	990	2
0.005	45	975	5
0.01	79	956	10
0.05	248	855	43
0.10	373	775	77
0.50	750	500	250
1	896	368	368
2	1000	250	500
3	1038	192	578
4	1055	157	630
5	1064	133	669
10	1071	77	775
50	1040	18	924
100	1027	9	955
500	1009	2	988

When $r = 0.50$ and $r = 2$, the numbers seem too clean to be coincidental; and when $r = 1$, $x^*/1000$ and $z^*/1000$ looks suspiciously like $1/e$. So we decided to examine a continuous model of our problem.

A Continuous Model and Solution

Letting

$$\Delta x_t = x_{t+1} - x_t, \qquad \Delta z_t = z_{t+1} - z_t, \qquad \text{and} \qquad \Delta t = t + 1 - t = 1,$$

our difference equations may be rewritten as

$$\frac{\Delta x_t}{\Delta t} = \frac{-r x_t}{r x_t + z_t}, \qquad \frac{\Delta z_t}{\Delta t} = \frac{r x_t - z_t}{r x_t + z_t}.$$

These suggest exploring the system of differential equations

$$\frac{dx}{dt} = \frac{-rx}{rx+z}, \tag{1}$$

$$\frac{dz}{dt} = \frac{rx-z}{rx+z}, \tag{2}$$

$$x(0) = 1, \qquad z(0) = 0.$$

Here we choose $x(0) = 1$ instead of $1{,}000$, so that variables like $z(t)$ can be interpreted as "amount of X-X-Y per initial amount of X-X."

We wish to determine the value of t, call it t^*, which maximizes $z(t)$. From the physical description of the problem, t^* lies strictly between 0 and 2 and at a critical point satisfying $z'(t^*) = 0$. Let x^* and z^* denote $x(t^*)$ and $z(t^*)$, respectively. Then, by (2), we have

$$rx^* = z^*. \tag{3}$$

Dividing equation (2) by equation (1) gives us

$$\frac{dz}{dx} - \frac{1}{rx} z = -1,$$

which is a linear differential equation in z. To solve it, we multiply by the integrating factor

$$\mu(x) = \exp\left(\int \frac{-1}{rx}\,dx\right) = \exp\left(-\frac{1}{r}\ln x\right) = x^{-1/r}$$

(since $x > 0$), giving

$$x^{-1/r}\frac{dz}{dx} - \frac{x^{-1/r}}{rx} z = -x^{-1/r},$$

$$\frac{d}{dx}\left(x^{-1/r} z\right) = -x^{-1/r}. \tag{4}$$

To integrate this equation with respect to x, we must distinguish the cases $r = 1$ and $r \neq 1$, as they correspond to different integration rules.

The Case $r = 1$

For the case $r = 1$, integrating both sides with respect to x gives us

$$z = -x \ln x + C.$$

Since at $t = 0$, we have $x(0) = 1$ and $z(0) = 0$, then $z(x) = 0$ when $x = 1$. Hence $C = 0$ and

$$z = -x \ln x.$$

At optimality, we have by equation (3), $z^* = x^*$. Plugging this expression into the above equation gives us $x^* = -x^* \ln x^*$. Hence,

$$z^* = x^* = 1/e,$$

as previously observed. As for t^*, we observe that, at all points in time, we have

$$t = 2 - 2x - z, \tag{5}$$

which can be derived from the equations $t = z + 2w$ (since the amount of Y equals the amount of X-X-Y plus twice the amount of Y-X-X-Y) and $x+z+w = x(0) = 1$ (since no X-X molecules are destroyed). Thus, we have

$$t^* = 2 - 2x^* - z^* = 2 - 3/e.$$

The Case $r \neq 1$

For the case $r \neq 1$, integrating both sides of (4) gives us

$$x^{-1/r}\, z = \frac{-r}{r-1}\, x^{(r-1)/r} + C,$$

i.e.,

$$z = \frac{-r}{r-1}\, x + C x^{1/r}.$$

Since $z(x) = 0$ when $x = 1$, we have $C = r/(r-1)$. Hence

$$z = \frac{r}{r-1}\left(x^{1/r} - x\right).$$

By (3), we can substitute $z^* = rx^*$, yielding

$$x^* = r^{r/(1-r)} \qquad \text{and} \qquad z^* = rx^* = r^{1/(1-r)}.$$

Hence, by (5), we have

$$t^* = 2 - 2x^* - z^* = 2 - 2r^{r/(1-r)} - r^{1/(1-r)}.$$

As a consistency check, we observe that

$$\lim_{r\to 0} t^* = 0, \qquad \lim_{r\to\infty} t^* = 1, \qquad \text{and} \qquad \lim_{r\to 1} t^* = 2 - 3/e;$$

and the formula

$$t^* = 2 - 2r^{r/(1-r)} - r^{1/(1-r)}$$

is consistent with the y^* values in **Table 1**.

The Method of Mass Action

To satisfy the chemist who may be suspicious about the equations that we derived using probability, we will now show that the same differential equations can be derived from the so-called "laws of mass action." We now consider all the molecules to be mixed together, and we describe the frequency of reactive collisions between molecules. The law merely states that the rate of any elementary reaction (that is, one that is not a sum of more fundamental reactions) is directly proportional to the concentrations of both reactants. To derive these differential equations from the laws of mass action, we proceed as follows. The chemist uses brackets to signify molar concentration, which has units of moles/liter.

Our reactions are as follows:

$$XX + Y \xrightarrow{k_1} XXY, \tag{6}$$

$$XXY + Y \xrightarrow{k_2} YXXY. \tag{7}$$

If we assume that these reactions are elementary and irreversible, we can write

$$\begin{aligned} \text{rate}_1 &= k_1 \cdot [XX] \cdot [Y], \\ \text{rate}_2 &= k_2 \cdot [XXY] \cdot [Y]. \end{aligned}$$

Conservation of mass enables us to determine a "rate equation" for each substance X-X, Y, X-X-Y, and Y-X-X-Y. We find the equation as follows. Since X-X-Y is created in (6) and destroyed in (7), the rate equation is $d[XXY]/dt = \text{rate}_1 - \text{rate}_2$. Thus

$$\begin{aligned} \frac{d[XX]}{dt} &= -k_1[XX][Y], \\ \frac{d[Y]}{dt} &= -k_1[XX][Y] - k_2[XXY][Y], \\ \frac{d[XXY]}{dt} &= k_1[XX][Y] - k_2[XXY][Y], \\ \frac{d[YXXY]}{dt} &= k_2[XXY][Y]. \end{aligned}$$

To simplify matters, it is convenient to normalize all the concentrations by dividing through by $[XX]_0$. Formally, we are making the substitutions

$$x = \frac{[XX]}{[XX]_0}, \qquad y = \frac{[Y]}{[XX]_0}, \qquad z = \frac{[XXY]}{[XX]_0}, \qquad \text{and} \qquad w = \frac{[YXXY]}{[XX]_0},$$

to generate the following equations:

$$\frac{dx}{dt} = -k_1'xy, \qquad \frac{dy}{dt} = -k_1'xy - k_2'zy,$$

$$\frac{dz}{dt} = k_1'xy - k_2'zy, \qquad \frac{dw}{dt} = k_2'zy,$$

where $k_1' = k_1[XX]_0$ and $k_2' = k_2[XX]_0$. The initial conditions in the new variables will be

$$x(0) = 1, \qquad z(0) = 0, \qquad w(0) = 0, \qquad y(0) = y_0,$$

where y_0 is unknown.

Our aim is to arrive at the same differential equations that we obtained using probability. To do this, we must realize that the time variable in the discrete model is not the same as the time variable in the mass action model. In the discrete model, one mole of y was added for each unit of time, so time could be used interchangeably with the amount of y that had reacted.

So, to generate the same differential equations, we shall choose an independent variable that represents the amount of y that has reacted. Like time, it increases as the reaction proceeds. This variable, call it ξ, will be given by $y_0 - y$: the original amount of y minus the amount present at the current time. We can think of ξ as a measure of the "extent of reaction."

Differentiating ξ with respect to time, we have

$$\frac{d\xi}{dt} = -\frac{dy}{dt} = k_1'xy + k_2'zy.$$

Eliminating time as an independent variable, we obtain

$$\frac{dx}{d\xi} = \frac{dx/dt}{d\xi/dt} = \frac{-k_1'xy}{k_1'xy + k_2'zy} = \frac{-k_1'x}{k_1'x + k_1'z},$$

$$\frac{dz}{d\xi} = \frac{dz/dt}{d\xi/dt} = \frac{k_1'xy - k_2'zy}{k_1'xy + k_2'zy} = \frac{k_1'x - k_2'z}{k_1'x + k_2'z},$$

$$\frac{dw}{d\xi} = \frac{dw/dt}{d\xi/dt} = \frac{k_2'w}{k_1'x + k_2'z}.$$

Since $dw/d\xi$ is not coupled with x and z (since Y-X-X-Y is an inert product), these equations can be restated as a system of two coupled nonlinear first order differential equations:

$$\frac{dx}{d\xi} = \frac{-rx}{rx + z}, \qquad \frac{dz}{d\xi} = \frac{rx - z}{rx + z},$$

with $x(0) = 1, z(0) = 0$ and, as before, $r = k_1/k_2 = k_1'/k_2'$. These are the same equations that were derived in the last section, based on a very different (and seemingly less realistic) model of molecular behavior. This derivation suggests that perhaps other problems in chemical kinetics might be attacked successfully by a probabilistic treatment.

References

Meehan, James, Edward Ferroni, and Laura Budrik. 1992. A mathematical solution to a chemical engineering problem. *The UMAP Journal* 13 (1): 7–16.

Yeremin, E.N. 1979. *The Foundations of Chemical Kinetics*. Moscow: MIR Publishers.

About the Authors

Arthur Benjamin is an assistant professor of mathematics at Harvey Mudd College and an associate editor of *The UMAP Journal*. His mathematical interests include combinatorics, probability, and operations research.

Gordon Hogenson is a chemistry major at Harvey Mudd College interested in mathematical modeling.