

ENTROPY IN BIOLOGY

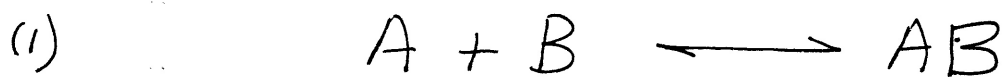
SPRING 2020

Charles S. Peskin

Lecture II : Bimolecular Reaction

- Microscopic equilibrium
- Markov chain consistent with microscopic equilibrium
- Application to event-driven simulation with fast and slow reactions
- Macroscopic chemical kinetics
- Diffusion-limited reaction rate

Thermodynamics of a microsystem
undergoing the reaction



in an incompressible ideal solution
at temperature T and pressure P

Let

$$(2) \quad n_0 = \# \text{ of molecules of solvent}$$

$$(3) \quad n_A = \# \text{ of molecules of } A + \# \text{ of molecules of } AB$$

$$(4) \quad n_B = \# \text{ of molecules of } B + \# \text{ of molecules of } AB$$

$$(5) \quad N_{AB} = \# \text{ of molecules of } AB$$

Then n_0, n_A, n_B are constant, and N_{AB} is a random variable with possible values

$$(6) \quad 0, 1, \dots, \min(n_A, n_B)$$

Also let

$$(7) \quad N_A = n_A - N_{AB} = \# \text{ of molecules of } A$$

$$(8) \quad N_B = n_B - N_{AB} = \# \text{ of molecules of } B$$

Although N_A and N_B are random variables, their values are determined by the value of N_{AB} .

The energy, volume, and entropy of the system are as follows

$$(9) \quad E = n_0 e_0(T) + N_A e_A(T) + N_B e_B(T) + N_{AB} e_{AB}(T)$$

$$(10) \quad V = n_0 v_0 + N_A v_A + N_B v_B + N_{AB} v_{AB}$$

$$(11) \quad S = n_0 \sigma_0(T) + N_A \sigma_A(T) + N_B \sigma_B(T) + N_{AB} \sigma_{AB}(T)$$

$$+ k \log M$$

where

$$\begin{aligned}
 (11.5) \quad M &= \frac{(n_0 + N_A + N_B + N_{AB})!}{(n_0)! (N_A)! (N_B)! (N_{AB})!} \\
 &= \frac{(n_0 + n_A + n_B - N_{AB})!}{(n_0)! (n_A - N_{AB})! (n_B - N_{AB})! (N_{AB})!}
 \end{aligned}$$

Note that $k \log M$ is the microscopic form of the entropy of mixing of the four species: solvent (0), A, B, AB when their numbers are n_0, N_A, N_B, N_{AB} , respectively, see Lecture 2.

Making use of (7-8), we can rewrite the expressions for E, V, S as follows

$$(12) \quad E = E_{N_{AB}=0} + N_{AB} \Delta e(T)$$

$$(13) \quad V = V_{N_{AB}=0} + N_{AB} \Delta v$$

$$(14) \quad S = S_{N_{AB}=0} + N_{AB} \Delta \sigma(T) + k \log M$$

where

$$(15) \quad \Delta e(T) = e_{AB}(T) - e_A(T) - e_B(T)$$

$$(16) \quad \Delta v = v_{AB} - v_A - v_B$$

$$(17) \quad \Delta \sigma(T) = \sigma_{AB}(T) - \sigma_A(T) - \sigma_B(T)$$

so that $\Delta e(T)$, Δv , $\Delta \sigma(T)$ are the changes in energy, volume and entropy (but not including the change in entropy of mixing) when one instance of the reaction $A + B \rightarrow AB$ occurs

The Gibbs free energy of the system is given by

$$(18) \quad G = E + PV - TS$$

Let $(\Delta G)_{n, n+1}$ be the change in G

that occurs when N_{AB} goes from $n \rightarrow n+1$.

This is possible for $n = 0, 1, \dots, \min(n_A, n_B) - 1$.

The change is to be made with T, P constant.

Let

$$(19) \quad M(n) = \frac{(n_0 + n_A + n_B - n)!}{(n_0)! (n_A - n)! (n_B - n)! n!}$$

Then

$$(20) \quad \frac{M(n+1)}{M(n)} = \frac{(n_A - n)(n_B - n)}{(n_0 + n_A + n_B - n)(n+1)}$$

Therefore

$$(21) \quad (\Delta G)_{n, n+1} = \Delta g(T, P) - kT \log \frac{(n_A - n)(n_B - n)}{(n_0 + n_A + n_B - n)(n+1)}$$

where

$$(22) \quad \Delta g(T, P) = \Delta e(T) - T \Delta \sigma(T) + P \Delta v$$

From (21),

$$(23) \quad e^{-\frac{(\Delta G)_{n,n+1}}{kT}} = \left(e^{-\frac{\Delta G(T,P)}{kT}} \right) \frac{(n_A - n)(n_B - n)}{(n_0 + n_A + n_B - n)(n + 1)}$$

In an equilibrium ensemble of our microsystems, let

$$(24) \quad p_n = (\text{probability that } N_{AB} = n)$$

Then

$$(25) \quad \frac{p_{n+1}}{p_n} = e^{-\frac{(\Delta G)_{n,n+1}}{kT}}$$

Now we combine (23) & (25), and also make the dilute solution approximation that there is much more solvent than any other species:

$$(26) \quad n_0 + n_A + n_B - n \approx n_0$$

This gives (with a ^{the parameter} defined below, see (30))

$$(27) \quad p_{n+1} = a \frac{(n_A - n)(n_B - n)}{(n+1)} p_n$$

and it follows by induction that

$$(28) \quad p_n = a^n \frac{(n_A)! (n_B)!}{(n_A - n)! (n_B - n)! (n)!} p_0$$

After multiplying and dividing by $n!$, we see that the combinatorial coefficient in (28) is equal to

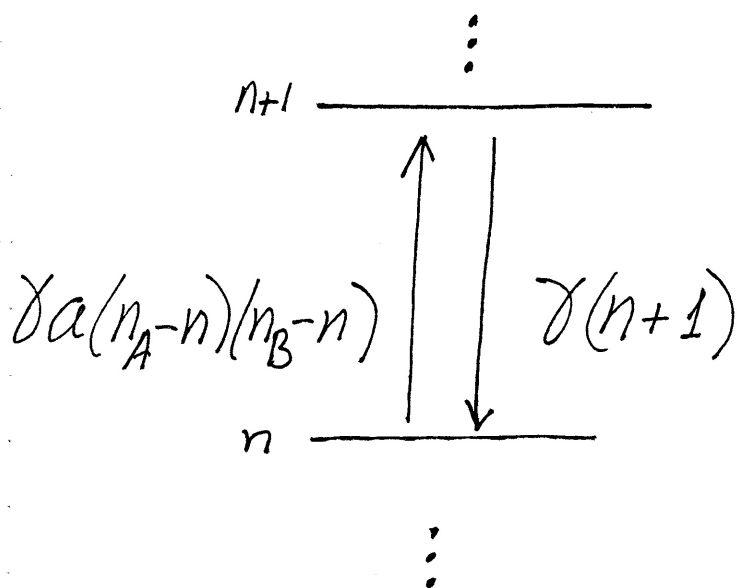
$$(29) \quad \binom{n_A}{n} \binom{n_B}{n} n!$$

and this is the number of different ways in which we can make n pairs of type AB when we have N_A items of type A and N_B items of type B from which to choose:

There are $\binom{N_A}{n}$ ways to choose n items of type A and $\binom{N_B}{n}$ ways to choose n items of type B, and once these two sets of n items each have been chosen, there are $n!$ ways to set up a 1-1 correspondence between them.

It is interesting that this combinatorial factor appears even though our entropy formula was based on the entropy of mixing without any reference to the chemical reaction.

It is now straight forward to set up a continuous-time Markov chain in which the states $n = 0, 1, \dots, \min(n_A, n_B)$ are the possible values of N_{AB} , and the transitions are steps of ± 1 .



The expressions on the arrows are the probabilities per unit time for the transitions (conditioned, of course, on the system being in the state at the tail of the arrow, so that the transition is possible)

The constant γ has units of $1/\text{time}$.

It may depend on temperature and pressure, but not on n , and not on any of the numbers n_0, n_A, n_B .

11

The parameter a that appears in the probability per unit time for the transition $n \rightarrow n+1$ is the same as the parameter a in equation (27). It is defined by

$$(30) \quad a = \frac{1}{n_0} e^{-\frac{\Delta g(T,P)}{kT}}$$

Even though thermodynamics only determines the ratio of the probabilities per unit time for the two transitions $n \rightarrow n+1$ and $n+1 \rightarrow n$, the way in which we have chosen the individual probability per unit time for each of the transitions is uniquely determined by the following considerations.

The transition $n+1 \rightarrow n$ involves the reaction $AB \rightarrow A+B$. In the state $n+1$ (which means $N_{AB} = n+1$) there are $n+1$ molecules of the type AB available to undergo this reaction.

These molecules are identical, so each of them has some probability per unit time γ of undergoing the dissociation reaction.

For the system as a whole, therefore, in the state $(n+1)$, the probability per unit time of a transition to the state n is $(n+1)\gamma$. As mentioned above, γ may depend on T, P , but it should not depend on any of the numbers n_0, n_A, n_B, n . With these additional assumptions together with the thermodynamic constraint, the probabilities per unit time are determined (except for their possible dependence on temperature and pressure).

Like the transition $n+1 \rightarrow n$, the transition $n \rightarrow n+1$ has a probability per unit time that is proportional to the number of entities available to make the transition. Such an entity is a pair of molecules, one of which is a free A and the other is a free B. The number of such pairs is $(n_A - n)(n_B - n)$. This number is multiplied by γ_a to get the probability per unit time of the transition. Note that a contains the factor $\frac{1}{n_0}$, see equation (30).

This is important, since $n_0 V_0$ is the volume of the system (in the dilute-solution limit). It is expected that a particular pair, one member of which is a free molecule of A and the other a free molecule of B, will take longer to find each other and react in a larger volume, but how much longer? We have here derived the result that the probability per unit time for such a pair to undergo $A + B \rightarrow AB$ is inversely proportional to the volume of the system, and therefore that the mean waiting time for K_{13} to occur is proportional to the volume of the system.

It is remarkable that we can make such a specific statement without a more detailed investigation into the manner in which A and B are moving and reacting.

Now consider the special case in which $n_A \gg n_B$. Then the possible values of n are $0 \dots n_B$, so $n \ll n_A$, and

$$(31) \quad \frac{n_A!}{(n_A - n)!} = n_A(n_A - 1) \dots n_A - n + 1 \approx n_A^n$$

In that case, our formula (28) for P_n becomes

$$(32) \quad P_n = (an_A)^n \binom{n_B}{n} P_0$$

and we have

$$(33) \quad 1 = \sum_{n=0}^{n_B} P_n = \left(\sum_{n=0}^{n_B} (an_A)^n \binom{n_B}{n} \right) P_0$$

$$= (1 + an_A)^{n_B} P_0$$

Making use of this in (32), we get

$$(34) \quad P_n = \frac{(an_A)^n}{(1+an_A)^{n_B}} \binom{n_B}{n}$$

$$= \left(\frac{an_A}{1+an_A} \right)^n \left(\frac{1}{1+an_A} \right)^{n_B-n} \binom{n_B}{n}$$

and this is just the binomial distribution probability

$$(35) \quad p = \frac{an_A}{1+an_A} = \frac{\frac{n_A}{n_0} e^{-\frac{\Delta g(T,P)}{kT}}}{1 + \frac{n_A}{n_0} e^{-\frac{\Delta g(T,P)}{kT}}}$$

$$(36) \quad 1-p = \frac{1}{1+an_A} = \frac{1}{1 + \frac{n_A}{n_0} e^{-\frac{\Delta g(T,P)}{kT}}}$$

Note that

$$(37) \quad \frac{n_A}{n_0} e^{-\frac{\Delta g(T,P)}{kT}} = \frac{n_A}{n_0 v_0} v_0 e^{-\frac{\Delta g(T,P)}{kT}}$$

$$= \frac{[A]}{K}$$

where

$$(38) \quad [A] = \frac{n_A}{n_0 v_0} = \text{concentration of } A$$

$$(39) \quad K = \frac{1}{v_0} e^{+\frac{\Delta g(T,P)}{kT}}$$

In (38) we are ignoring the distinction between $[A]$ and $[A] + [AB]$, and this is justified because $n_A \gg n_B$. Also, we are regarding $n_0 v_0$ as the total volume of the system, as it essentially is in the dilute-solution limit.

The constant K is the equilibrium constant (dissociation constant) of the reaction.

Note that K has units of concentration. In terms of K ,

$$(40) \quad p = \frac{[A]}{K + [A]}, \quad 1-p = \frac{K}{K + [A]}$$

What we have found is that when there are many more molecules of A than of B in the system, each molecule of B is functioning independently of the others, and for any one molecule of B , the probability that it will be found with a molecule of A bound to it (or the fraction of time that it will have a molecule of A bound to it) is given by the above expression for p .

Note that it is perfectly possible for $n_A \gg n_B$ so that the above is applicable, and yet for p to be small so that most molecules of B are not bound to A . This simply requires

$[A] \ll K$, which is equivalent to $\frac{n_A}{n_0} \ll e^{-\frac{\Delta G(T,P)}{kT}}$

This will always be true when the solution is sufficiently dilute, and it is not at all inconsistent with $n_A \gg n_B$.

From (30) & (39) we see that

(41) $aK = \frac{1}{n_0 v_0} = \frac{1}{V}$

where V is the volume of the system.

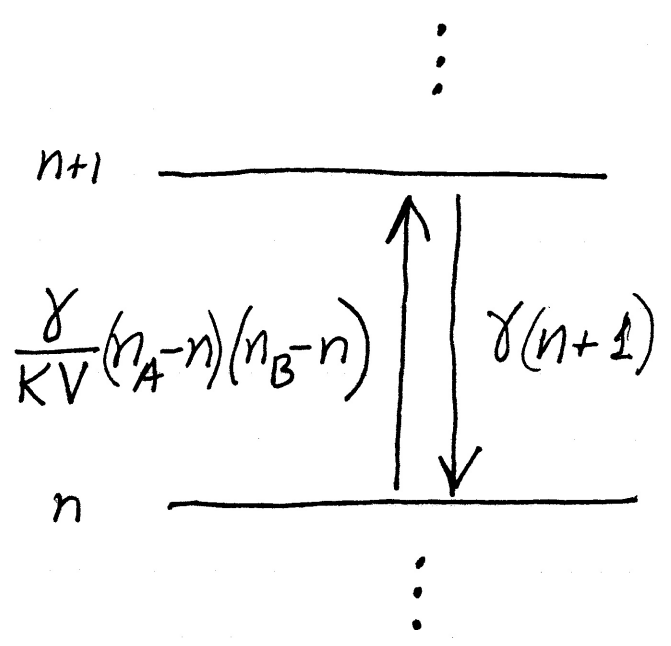
This is useful because it lets us rewrite some key results in terms of macroscopic quantities, namely the equilibrium constant and the volume of the system

In particular, equations (27-28) become

$$(42) \quad P_{n+1} = \frac{1}{KV} \frac{(n_A - n)(n_B - n)}{(n+1)} P_n$$

$$(43) \quad P_n = \left(\frac{1}{KV}\right)^n \frac{(n_A)! (n_B)!}{(n_A - n)! (n_B - n)! (n)!} P_0$$

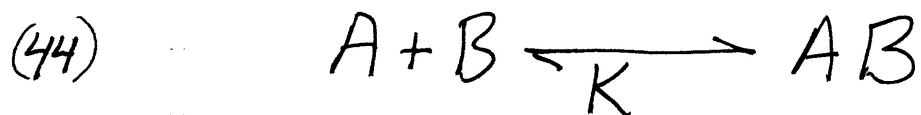
and the Markov chain of the microscopic system now looks like this :



Here we see explicitly the influence of volume on the probability per unit time of $A+B \rightarrow AB$

As an application of the foregoing analysis of microscopic equilibrium, we show how it can be used in event-driven simulation of a system with a mixture of fast and slow reactions.

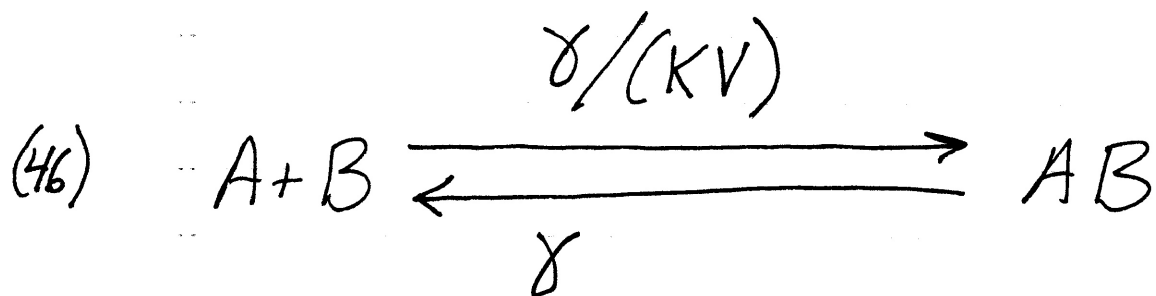
Consider the reaction scheme



Thus B is an enzyme which is degrading A . The number of molecules of B in the system is conserved, but the number of molecules of A is decreasing.

We would like to simulate this system by event-driven simulation.

The standard way to do this would be to assign rate constants to the first reaction, which we know how to do:



In the case of $\gamma/(kV)$ this is the microscopic rate constant. It multiplies the number of pairs consisting of a free A and a free B to get the probability per unit time of an occurrence of $A + B \rightarrow AB$.

The corresponding macroscopic rate constant will be discussed later.

For a first-order reaction like $AB \rightarrow A + B$, there is no distinction between the microscopic and the macroscopic rate constant.

Now we have three possible reactions, each with a specified probability per unit time that depends on the state of the system, and each with a specified change in state:

reaction	prob/time	change in		
		N_A	N_B	N_{AB}
$A+B \rightarrow AB$	$\frac{\gamma}{KV} N_A N_B$	-1	-1	+1
$AB \rightarrow A+B$	γN_{AB}	+1	+1	-1
$AB \rightarrow B$	αN_{AB}	0	+1	-1

This simulation can certainly be done, but what if we are interested in the limit $\gamma \rightarrow \infty$? The larger we make γ , the more events we get corresponding to the first two reactions, and so the computation eventually bogs down.

We can avoid the above difficulty by treating the reaction (44) as a reaction in equilibrium, even though the system as a whole is not at equilibrium. The state variables are now

$$(47) \quad n_A = N_A + N_{AB}$$

$$(48) \quad n_B = N_B + N_{AB}$$

The value of n_B is constant, and the value of n_A changes in steps of (-1) whenever the reaction $AB \rightarrow B$ occurs. These are the only events that we consider. The probability per unit time for the occurrence of such an event is

$$(49) \quad \propto E[N_{AB} | n_A, n_B]$$

$$= \alpha \sum_{n=0}^{\min(n_A, n_B)} n p_n$$

with P_n given by the recursion relation (42) or by the more explicit formula (43), and in either case with the normalization $\sum P_n = 1$. We should not assume

here that $n_A \Rightarrow n_B$, since n_A is decreasing to zero, and at some point it will become comparable to and then less than n_B . Indeed, one goal of the simulation might be to sample the distribution of times required for the complete elimination of A from the system.

Note that there are at worst n_B terms in the sum on the right-hand side of (49), and n_B is constant.

The most efficient way to evaluate P_n is to use the recursion relation (42) starting from a provisional value $P_0 = 1$, and then to normalize the result by dividing by $\sum P_n$.

Macroscopic Kinetics of $A+B \rightleftharpoons AB$

The equations of macroscopic chemical kinetics can be written down immediately by inspection from our microscopic scheme. We do this in two steps: first using the microscopic variables as if they were real instead of integer-valued, and then rewriting the equations in terms of concentrations.

In terms of the variables N_A , N_B , and N_{AB} , we have

(50)
$$\frac{dN_A}{dt} = -\frac{\gamma}{KV} N_A N_B + \gamma N_{AB}$$

(51)
$$\frac{dN_B}{dt} = -\frac{\gamma}{KV} N_A N_B + \gamma N_{AB}$$

(52)
$$\frac{dN_{AB}}{dt} = \frac{\gamma}{KV} N_A N_B - \gamma N_{AB}$$

Note that

$$(53) \quad \frac{d}{dt} (N_A + N_{AB}) = 0$$

$$(54) \quad \frac{d}{dt} (N_B + N_{AB}) = 0$$

Thus, there are two conserved quantities

$$(55) \quad n_A = N_A + N_{AB}$$

$$(56) \quad n_B = N_B + N_{AB}$$

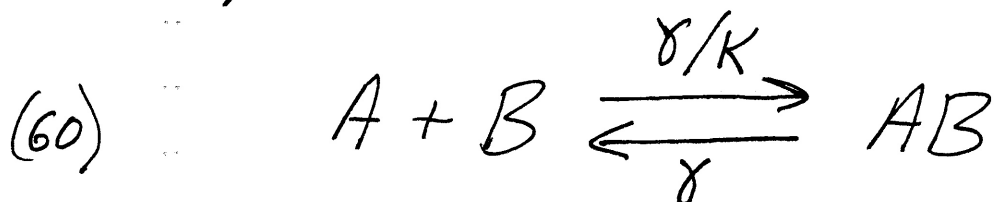
We can rewrite (50-52) in terms of concentrations by dividing both sides of each equation by V . Note that the extra factor of V that we need for the nonlinear terms is already conveniently there. In this way we get

$$(57) \quad \frac{d[A]}{dt} = -\frac{\gamma}{K} [A][B] + \gamma [AB]$$

$$(58) \quad \frac{d[B]}{dt} = -\frac{\gamma}{K} [A][B] + \gamma [AB]$$

$$(59) \quad \frac{d[AB]}{dt} = \frac{\gamma}{K} [A][B] - \gamma [AB]$$

The diagram corresponding to these equations is



Comparison with (46) shows that the macroscopic rate constant for the bimolecular reaction $A+B \rightarrow AB$ is not the same as the microscopic rate constant. Microscopic rate constants always have units of $1/\text{time}$, but the macroscopic rate constant here has units of $\text{volume}/\text{time}$.

Another important remark is that the macroscopic equations obscure the influence of volume on the rate of a bimolecular reaction. That influence is still present, although hidden.

If we start with an equilibrium state of the above system and suddenly double the volume by adding pure solvent to the system, all of the concentrations will suddenly be cut in half, and the term $[A][B]$ is thereby reduced by a factor of 4 whereas $[AB]$ is only reduced by a factor of 2. Thus the system will no longer be in equilibrium and changes in concentration will follow until a new equilibrium is established.

An equilibrium of equations (57-59) satisfies the following equations:

$$(61) \quad [A][B] = K [AB]$$

$$(62) \quad [A] + [AB] = [A]_0$$

$$(63) \quad [B] + [AB] = [B]_0$$

Here $[A]_0$ and $[B]_0$ are the conserved quantities corresponding to n_A and n_B .

Specifically

$$(64) \quad [A]_0 = n_A/V, \quad [B]_0 = n_B/V$$

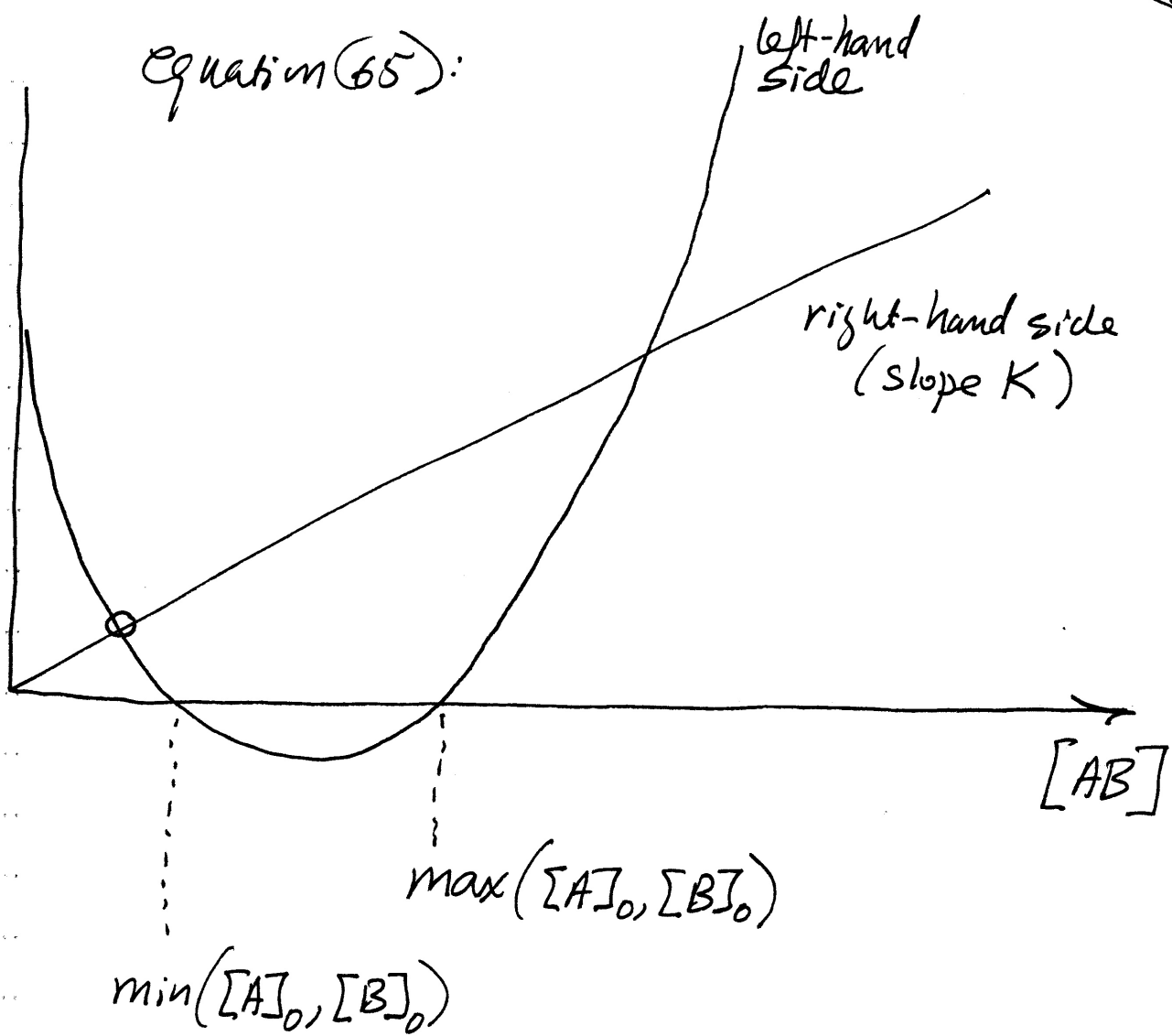
Equations (61)-(63) can be rewritten as a single equation for $[AB]$:

$$(65) \quad ([A]_0 - [AB])([B]_0 - [AB]) = K [AB]$$

By plotting the two sides of (65), we see immediately that there are two positive solutions, but only the smaller of them lies in the interval

$$(66) \quad (0, \min [A]_0, [B]_0)$$

as required so that $[A]$ and $[B]$ will be positive.



The acceptable solution is

$$(66.5) \quad [AB] = \frac{([A]_0 + [B]_0 + K) - \sqrt{([A]_0 + [B]_0 + K)^2 - 4[A]_0[B]_0}}{2}$$

As the plot shows, $[AB] \rightarrow \min([A]_0, [B]_0)$ as $K \rightarrow 0$. How it achieves this gymnastic feat is interesting!

An interesting exercise related to equation (65) and its solution (66.5) is to evaluate

(67) $\lim_{[B]_0 \rightarrow 0} \left(\frac{[AB]}{[B]_0} \right)$

This is related to the analysis of the microscopic system when $n_A \gg n_B$.

It would also be interesting to compare the macroscopic equilibrium concentration $[AB]$ to the computed value of

(68) $\frac{1}{V} E [N_{AB} | n_A, n_B]$

See (49), with n_A and n_B related to $[A]_0$ and $[B]_0$ by equation (64).

The results should be close, but not (I think) the same. They should converge, however, as V is scaled up with $[A]_0$ and $[B]_0$ held constant.

Diffusion limited rate of a bimolecular reaction

A striking feature of the foregoing is that the microscopic rate constant for $A+B \rightarrow AB$ is inversely proportional to the volume of the system. It is only because of this inverse proportionality that we are able to write chemical kinetic equations for the macroscopic concentrations that do not explicitly involve the volume of the system.

The purpose of this section is to see whether we can arrive at a mechanistic understanding of this phenomenon.

We begin with a macroscopic, ^{steady-state} diffusion problem. The domain is the exterior of a sphere of radius R in \mathbb{R}^3 . On this domain, we solve

$$(69) \quad \nabla \cdot \underline{F} = 0$$

$$(70) \quad \underline{F} = -D \nabla c$$

in spherical polar coordinates, with boundary condition

(71) $c(R) = 0$

(72) $c(\infty) = c_\infty$

The solution is expected to be of the form

(73) $\underline{f} = \frac{-F}{4\pi r^2} \underline{\hat{r}}$

where $\underline{\hat{r}}$ is the unit vector in the radial direction. This satisfies $\nabla \cdot \underline{f} = 0$.

Substituting (73) into (70), we get

(74) $\frac{F}{4\pi r^2} = D \frac{\partial c}{\partial r}$

Since $c(R) = 0$, this implies

(75) $c(r) = \frac{F/D}{4\pi R} - \frac{F/D}{4\pi r}$

Then, letting $r \rightarrow \infty$ and using the

boundary condition (72), we get

$$(76) \quad F = 4\pi DR C_{\infty}$$

Note that D has units of length²/time, and of course R has units of length and C_{∞} has units of 1/volume, so F has units of 1/time.

We interpret this result as follows:

If a sphere of radius R with an absorbing boundary is immersed in a sea of particles with number per unit volume C_{∞} , and if each particle has diffusion coefficient D , then the sphere will be hit by particles at the rate F .

Now consider such a sphere in a finite volume $V \gg \frac{4}{3}\pi R^3$, and

suppose there is one particle, placed inside that volume at a random location at $t=0$, and we want to know the probability per unit time that the particle will encounter the with diffusion coefficient D

sphere (given that it has not yet done so). Since the concentration of particles is now $1/V$, this probability per unit time should be given, at least approximately, by the above formula for F with $c_{\infty} = 1/V$, that is

$$(77) \quad F = \frac{4\pi DR}{V}$$

(It may be argued that the particle did not start "at ∞ ", but since $V \gg \frac{4}{3}\pi R^3$, it may as well

have. Equation (77) should be regarded as ^{an} asymptotic result, valid as $R \rightarrow 0$.)

Now consider a pair of particles with radii r_A and r_B . For simplicity, assume that particle B is not moving and that particle A has diffusion coefficient D_A . Since the particles touch when their centers come within $r_A + r_B$ of each other, the probability per unit time of contact,

given that contact has not already occurred,
is given by

$$(78) \quad F = \frac{4\pi D_A (r_A + r_B)}{V}$$

Note, however that D_A is determined by r_A . The diffusion coefficient of a sphere of radius r_A in a fluid of density ρ and kinematic viscosity ν is given by

$$(79) \quad D_A = \frac{kT}{6\pi\rho\nu r_A}$$

Substituting (79) into (78), we get

$$(80) \quad F = \frac{2}{3} \frac{kT}{\rho V \nu} \left(1 + \frac{r_B}{r_A} \right)$$

For example, if $r_A = r_B$

$$(81) \quad F = \frac{4}{3} \frac{kT}{\rho V \nu}$$

Note that these results scale with V in the right way.

Some approximate numbers with V as the volume of a red blood cell and with ρ, ν as the density and kinematic viscosity of water are

$$(82) \quad kT = 4 \times 10^{-21} \text{ Kg} \cdot \text{m}^2/\text{s}^2$$

$$(83) \quad \rho = 10^3 \text{ Kg}/\text{m}^3$$

$$(84) \quad V = 10^{-16} \text{ m}^3$$

$$(85) \quad \nu = 10^{-6} \text{ m}^2/\text{s}$$

and then

$$(86) \quad F = \frac{4}{3} \frac{4 \times 10^{-21} \text{ Kg} \cdot \text{m}^2/\text{s}^2}{(10^3 \text{ Kg}/\text{m}^3)(10^{-16} \text{ m}^3)(10^{-6} \text{ m}^2/\text{s})}$$

$$\cong 5 \times 10^{-2}/\text{s} = 1/(20 \text{ s})$$

From equation (80), we see that the rate F can be increased above that given by (81) by making the fixed target molecule B have a larger radius than the mobile molecule A (and in that case, it is all the more reasonable to model B as fixed, since a larger particle even if free has a smaller diffusion coefficient).

In summary, we have seen that when the rate of the reaction $A + B \rightarrow AB$ is estimated by considering diffusion with reaction in contact, the estimated probability per unit time scales with system volume in the manner predicted by thermodynamics, and also that the reaction rate has a finite limit as the particle sizes approach zero in a fixed ratio, because faster diffusion compensates for the greater difficulty in finding a smaller target.