

Process model formulation and solution, 3E4

Assignment 1

Kevin Dunn, dunnkg@mcmaster.ca

September 2010

The solutions were prepared by both TA's: Ali and Elliot.

Question 1 [1]

Note: Distillation tray modelling.

The distillation column is a common unit operation in refineries, air-separation, and fine-chemical processes. The figure, taken from Wikipedia, shows a cross-section of two trays in a distillation column.

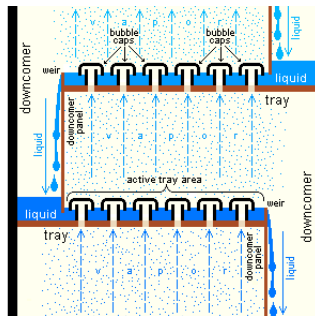
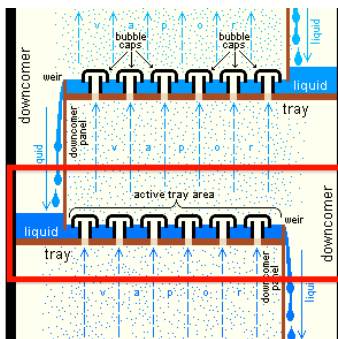


Figure 1: Source: http://en.wikipedia.org/wiki/File:Bubble_Cap_Trays.PNG

1. Redraw the figure, and add to it a boundary you would use to write the mass and energy balance equations for a single tray.
2. Which physical phenomena would you need to consider for an accurate model?
3. Which reasonable simplifying assumptions would you make to reduce the complexity of your model?

Solution

1. The following boundary would be a good basis for energy and mass balances on a single tray:



2. In order to construct the mass balances you would need to account for:

- Liquid mass flow (both bulk and of $n - 1$ individual components) into the tray's downcomer from the tray above.
- Liquid mass flow (both bulk and of $n - 1$ individual components) out of the tray into the next tray's downcomer.
- Vapour mass flow (both bulk and of $n - 1$ individual components) into the bottom of the tray from the tray below (condensation through the bubble caps).
- Vapour mass flow (both bulk and of $n - 1$ individual components) off the top of the tray (evaporation).
- Either a form of thermodynamic equilibrium calculation to calculate the steady state vapour to liquid ratios or a form of calculation (likely empirical) to calculate the dynamic vapour flow rates from each tray.
- In the more complicated case, thermodynamic relations to calculate mixture properties within each tray.

In order to construct an energy balance you would need to account for:

- Convective flow of energy in via liquid flow into the tray's downcomer from the tray above.
- Convective flow of energy out via liquid flow out of the tray into the next tray's downcomer.
- Convective flow of energy in via vapour flow into the bottom of the tray from the tray below (condensation through the bubble caps).
- Convective flow of energy out via vapour flow off the top of the tray (evaporation).
- Thermodynamic relations to calculate the relative energy of the liquid and vapour mixtures at the internal conditions.
- Depending on how complicated the model is desired to be one could also account for convective heat loss off of the sides of the column.
- Radiative transfer between trays.

3. In order to simplify the model one could make any of the following assumptions (whether these assumptions are reasonable depends on the system being modelled):

- Perfect mixing of liquid resulting in uniform temperature and concentration at each tray.
- 100% condensation of vapour entering the bottom of each tray.
- All vapour leaving a tray enters the next tray (i.e. no condensation and backflow in the vapour transfer area).
- All vapour transmission occurs over the active tray area.
- No vapour transfer within the downcomer regions.
- The vapour and liquid of each tray are at equilibrium.
- Perfectly insulated column (i.e. no radiative losses).
- Negligible radiative transfer between trays.
- The components and mixtures are non-reactive.
- The effects of mixing are negligible.

Question 2 [2]

Note: This question will slowly build up the equations and assumptions for solving a common chemical engineering problem: heat loss from a pipe.

Our final goal is to be able to predict the temperature profile, along the length of a tube. Such tubes could occur, for example, in a shell and tube heat exchanger, as shown in the figure below from the Wikipedia article.

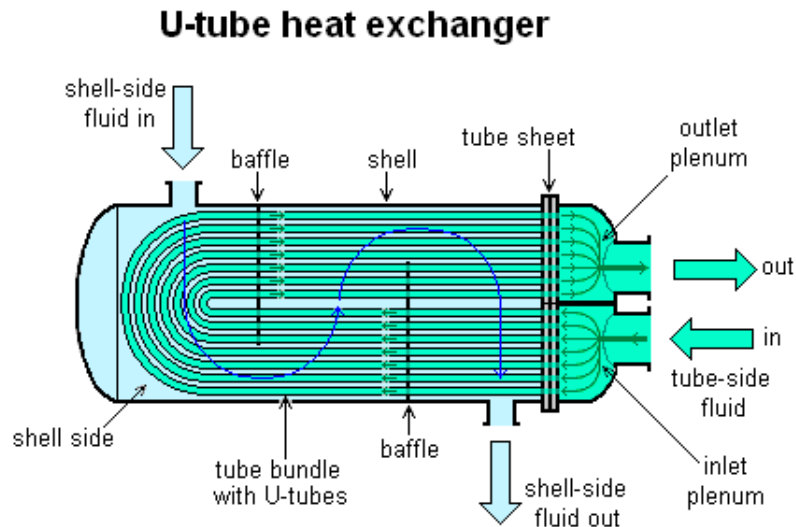
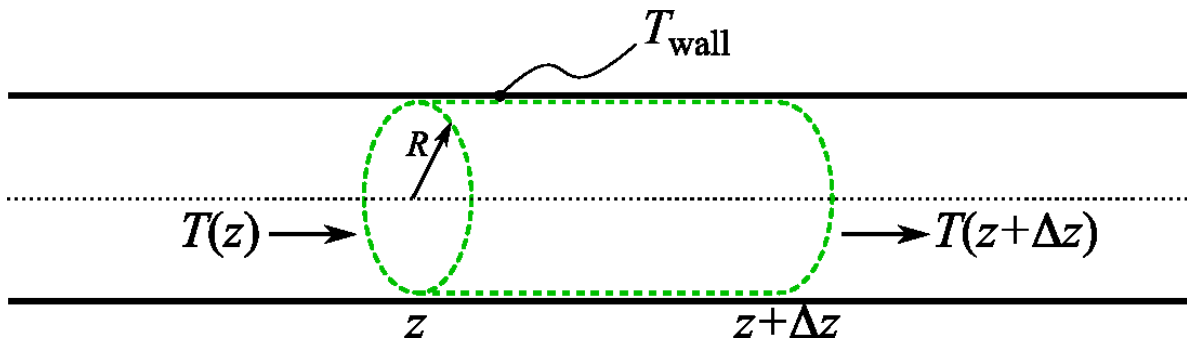


Figure 2: Source: http://en.wikipedia.org/wiki/Shell_and_tube_heat_exchanger

1. Write down all the possible controlling mechanisms you might consider for a high-accuracy model of this shell and tube heat exchanger.
2. Now let's start simple, and consider only the fluid flow in the pipe. What would the cross-sectional profile for the flow look like under (a) laminar flow and (b) turbulent flow?
3. Consider a short cross-sectional segment, of width Δz , where the temperature is $T(z)$ at the entry and $T(z+\Delta z)$ at the exit. The figure below contains the boundary region already drawn, in green.



- (a) Which specific assumption do you have to invoke so that the temperature in the radial direction is constant?
- (b) Assuming this assumption is true, write down the average temperature within the segment, and define it as $\bar{T}(z)$.
- (c) What is the surface area of the pipe in this segment?

- (d) Finally, using Newton's law of cooling, write down the rate of energy loss, Q , to the outside environment, where the wall temperature is defined as T_{wall} . What are the units of the heat transfer coefficient, and the units of Q ?
4. Now perform an energy balance over this segment.
- Heat in =
 - Heat out =
 - Heat generated =
 - Overall heat balance, assuming steady-state:
5. Lastly, divide this steady-state equation by Δz and take limits so that $\Delta z \rightarrow 0$. This allows you to reach our goal and write an equation that predicts the temperature profile along the length of a tube.

Recall from calculus that:

$$\lim_{\Delta z \rightarrow 0} \frac{T(z + \Delta z) - T}{\Delta z} = \frac{dT}{dz}$$

Solution

1. A shell and tube heat exchanger operates by transferring heat from the hot stream to the cold stream through conduction. Therefore one controlling mechanism is *heat conduction*. Other controlling mechanisms include:
 - Fluid flow: which exists in the tube and shell sides
 - Evaporation: Some heat exchangers involve phase changes. For example, the cold fluid evaporates (e.g. reboilers), or a vapour fluid condenses (e.g. the condenser in a refrigerator cycle). In such cases, the phase change must be considered.
 - Reaction: Many industrial packed-bed reactors have a shell and tube design where the reaction occurs in the tube side. The shell side acts as a cooling or heating medium by transferring heat from or to the tube side (reaction medium).
 - Radiation: Heat transfer via thermal radiation from the body of the heat exchanger to ambient or *vice versa*.
2. In laminar flow, the cross-sectional profile for the velocity is parabolic where the maximum velocity occurs in the centre of the pipe. Laminar flow occurs at low Reynolds numbers ($Re < 2100$ for pipes). In turbulent flow, however, the radial distribution of velocity is noticeable only within a thin layer adjacent to the wall. Turbulent flow occurs at high Reynolds numbers ($Re > 2100$ for pipes).
3. Considering a segment of width Δz :
 - (a) We assume that we have a *plug flow* through the pipe, which means that the velocity is constant in the radial direction. Note that the plug flow assumption requires a turbulent flow where the boundary layer adjacent to the wall of the pipe is negligible.
 - (b) The average temperature can be obtained from the temperatures at the two ends of the segment as $\bar{T} = \frac{T(z) + T(z + \Delta z)}{2}$
 - (c) The surface area of the cylindrical segment is $A = 2\pi R\Delta z = \pi D\Delta z$ ($D = 2R$).
 - (d) Newton's law of cooling states that the rate of the heat transferred through the wall (Q) is proportional to the difference between the fluid's temperature and the wall temperature. We choose the proportional constant as $(\pi D\Delta z)h$ where h is positive and is the heat transfer coefficient. Therefore, $Q = (\pi D\Delta z)h(\bar{T} - T_{wall})$, where \bar{T} is the fluid's temperature inside the segment. The unit of Q is energy per time (e.g. $J s^{-1}$). As a result, the unit of the constant h will be $J s^{-1} m^{-2} K$.

4. The energy balance at steady state is

$$0 = \dot{m}c_p (T(z) - T(z + \Delta z)) - Q$$

where \dot{m} is the mass flowrate inside the pipe, and Q is the heat transferred through the wall. Substituting Q from Newton's law gives $0 = \dot{m}c_p(T(z) - T(z + \Delta z)) - (\pi D \Delta z)h(\bar{T} - T_{wall})$.

Then rearranging, and dividing both sides by $\dot{m}c_p \Delta z$ yields $\frac{T(z + \Delta z) - T(z)}{\Delta z} = -\frac{(\pi D)h}{\dot{m}c_p}(\bar{T} - T_{wall})$.

The term $\frac{T(z + \Delta z) - T(z)}{\Delta z}$ will represent the derivative of T with respect to z as $\Delta z \rightarrow 0$. Also, $T = \bar{T}$ when $\Delta z \rightarrow 0$. Let us define $\lambda = \frac{(\pi D)h}{\dot{m}c_p}$. Therefore, the following equation results:

$$\frac{dT}{dz} = -\lambda(T - T_{wall})$$

This is a differential equation that predicts the fluid's temperature versus the length of the pipe, at steady state.

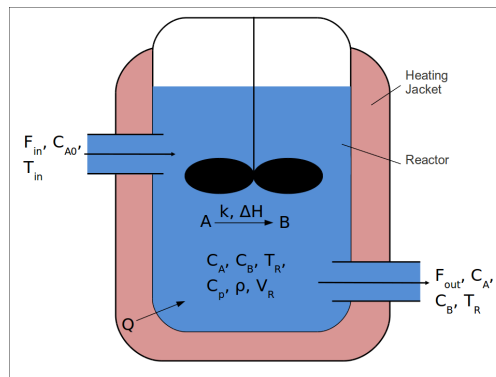
Question 3 [1.5]

Consider a CSTR where an irreversible, first-order endothermic reaction of the form $A \xrightarrow{k} B$ takes place. Let C_A denote the concentration of the species A in the reactor, T_R and T_{in} denote the temperatures of the reactor and of the inlet stream, respectively, Q , is the heat added to/removed from the reactor, C_{A0} is the concentration of A in the inlet stream, V is the volume of the reactor, k_0 , E , ΔH are the pre-exponential constant, the activation energy, and the enthalpy of the reaction and C_p and ρ are the heat capacity and fluid density in the reactor.

Develop a model that describes the evolution of the concentration and temperature in the reactor, using a systematic modelling approach that also outlines all assumptions made.

Solution

We start by generating a diagram of the system to aid in visualizing the set up and constructing the mass and energy balances:



In order to simplify the modelling process we make the following reasonable assumptions (based on the implications of the problem statement):

- Either the reagents A and B have similar heat capacities and densities/ mixture densities or the reagents are in dilute enough concentration that the heat capacity and density of the carrier fluid is dominant (i.e. C_p and ρ are constant)

- The reactor is perfectly mixed
- Shaft work is negligible
- Kinetic and potential energies are negligible
- The reactor is sealed (i.e. convective and evaporative losses are negligible)
- All physical properties are temperature independent
- Effects of mixing may be neglected (i.e. ideal mixtures)
- The system has a constant volume (i.e. $F_{in} = F_{out} = F$)
- There is no reagent B in the inlet

Moving to the modelling of the concentration of A and B in the reactor. It is important to note that a proper mass/molar balance deals with absolute quantities, not concentration. The only reason we can separate C_a/C_b and V_R in such a nice way is because we assumed a constant reactor volume. In the more complicated case we would need to consider the chain rule, which would result in a $\frac{dV_R}{dt}$ term (In actuality we have used the chain rule below but since V_R is assumed constant the term $\frac{dV_R}{dt}$ simply cancels to 0).

$$\frac{dn_a}{dt} = \frac{dC_a V_R}{dt} = \frac{dC_a}{dt} V_R = \underbrace{\sum_i (C_{a0,i} \times F_{in,i})}_{\text{inlet}} - \underbrace{\sum_i (C_a \times F_{out,i})}_{\text{outlet}} + \text{Generation} - \text{Consumption}$$

$$\frac{dn_b}{dt} = \frac{dC_b V_R}{dt} = \frac{dC_b}{dt} V_R = \underbrace{\sum_i (C_{b0,i} \times F_{in,i})}_{\text{inlet}} - \underbrace{\sum_i (C_b \times F_{out,i})}_{\text{outlet}} + \text{Generation} - \text{Consumption}$$

We know that there is no generation of A and no consumption of B in the reactor and we know there is only one inlet and outlet stream. Therefore:

$$\begin{aligned} \frac{dC_a}{dt} &= C_{a0} \frac{F}{V_R} - C_a \frac{F}{V_R} - k_0 e^{-\frac{E}{RT_R}} C_a \\ \frac{dC_b}{dt} &= -C_b \frac{F}{V_R} - k_0 e^{-\frac{E}{RT_R}} C_a \end{aligned}$$

Moving to the energy balance (we again take note that the only reason the $\frac{dE}{dt}$ may be whittled down to $\frac{dT}{dt}$ is because all other terms are assumed constant). Note that “other sources” refers to things such as shaft work, kinetic/potential energy, etc.:

$$\frac{dE}{dt} = \sum_i (\text{Convective flow in}) - \sum_i (\text{Convective flow out}) - \text{Heat of reaction} + Q + \text{other sources}$$

$$\begin{aligned} \frac{d(C_p \rho V_R (T_R - T_{ref}))}{dt} &= \frac{d(C_p \rho V_R T_R)}{dt} - \frac{d(C_p \rho V_R T_{ref})}{dt} = \frac{d(C_p \rho V_R T_R)}{dt} - (0) \\ \frac{d(C_p \rho V_R T_R)}{dt} &= (C_p \rho F (T_{in} - T_{ref})) - (C_p \rho F (T_R - T_{ref})) - \Delta H k_0 V_R e^{-\frac{E}{RT_R}} C_a + Q + (0) \end{aligned}$$

Simplifying:

$$\frac{dT_R}{dt} = \left(\frac{F}{V_R} \right) (T_{in} - T_R) - \left(\frac{\Delta H k_0}{C_p \rho} \right) C_a e^{-\frac{E}{RT_R}} + \frac{Q}{C_p \rho V_R}$$

It is very important to note that T_{ref} was used to calculate each enthalpy term. In this specific case the T_{ref} cancels out due to the assumption made at the start of the solution (specifically the assumptions of equal inlet and outlet flow rates and an equal and constant density and heat capacity over the entire mixture). A reference temperature **must always be used when calculating enthalpies. There is no such thing as an absolute enthalpy.** As the current state of thermodynamics stands it is theoretically impossible to calculate absolute enthalpy (who knows what the future will bring...). As standard practice you should always start with T_{ref} in your energy balances and if it cancels out then all the better for your solution.

Question 4 [1.5]

The simplest example of biological wastewater treatment involves removing a single nutrient, N, from a liquid waste stream and converting it to (solid) biomass, B. This process is used, for example, by pop manufacturers to treat liquid waste from their processes, where N would be sugar (e.g. fructose). The microorganisms absorb N from the wastewater and use it to grow more and more biomass.

1. Draw a simple CSTR where this liquid stream enters, containing N_{in} [mg/L] and the stream leaving the reactor contains both liquid and solid fractions. The solids in the outlet stream can be separated, and the remaining liquid would contain much less N, allowing the company to discharge that liquid to municipal wastewater treatment systems.

Draw the boundaries you will consider on your diagram as well.

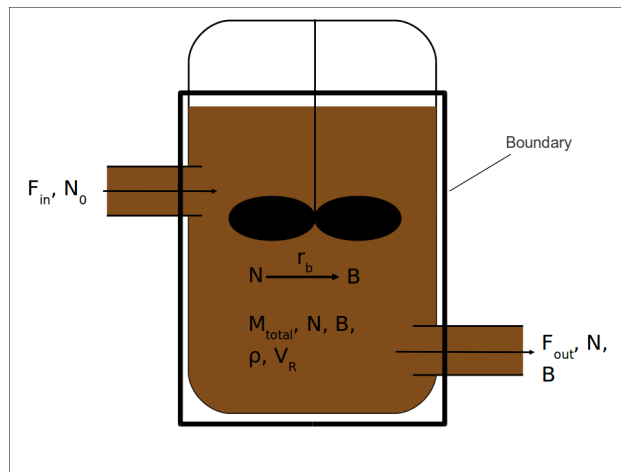
2. Write an overall mass balance, where the flows, q , are expressed in units of litres per day, typically being in the order of megalitres per day.
3. Write a dynamic mass balance for N and B, where their units are expressed in mg/L. Assume there is no biomass in the inlet stream. The reaction rate for converting $N \rightarrow B$ is given by r_B in units of mg/(L.day).
4. Write the two equations for the steady-state values of N and B if $r_B = \mu_{max} \frac{N}{K + N} B$, where K [mg/L] and μ_{max} [1/day] are constants from the Monod rate expression. Note that $r_N = -\frac{1}{Y_B} r_B$, where Y_B is the yield of B from N.
5. Are these equations solvable by hand if you are given the constants such as N_{in} , K , μ_{max} , etc in your equations?

Solution

1. A diagram of the process is provided below. The boundary that will be used for the component and overall mass balances has also been depicted and labelled (essentially the system will be the liquid contents of the reactor).

In order to simplify the modelling process we make the following reasonable assumptions:

- The overall density of the mixture is roughly constant and equal to that of water (This may seem like a huge leap at first but you must consider the fact that N is likely dilute, otherwise the company would try to reclaim these resources, and the organisms being grown are mostly water and so the density of the overall mixture will likely be close to that of water, the carrier fluid). In practice this assumption leads to reasonably good results.
- The reactor is perfectly mixed.



- The reactor is sealed (i.e. convective and evaporative losses are negligible). This assumption has a few caveats. You cannot assume this for general biological systems and in reality even this system will have gaseous losses but we are not modelling these components. Waste decomposition processes (ex. activated sludge, anaerobic digestion) will almost always convert at least a portion of N to some gaseous component (ex. CO_2 , CH_4 , etc.). But since we are not modelling these components we can simply forget about gas loss terms. It is also possible that in a real system evaporative losses of water would occur over time but these would likely be negligible in comparison to the relative size of the reactors.
 - All physical properties are temperature independent.
 - Effects of mixing may be neglected (i.e. ideal mixtures).
 - The system has a constant volume (i.e. $F_{in} = F_{out} = F$).
 - The inlet stream is sterile (i.e. no active organisms). Again this assumption is context sensitive. In activated sludge plants, for example, there will almost always be a recycle stream to reuse organisms (“bugs” are resources too). For the purposes of this question we will assume no recycling and no organisms in the inlet stream in general (the waste stream could be contaminated...)
 - Endogenous decay (i.e. bacteria death) is negligible.
2. In writing the overall mass balance we recall that volume **is not a conserved quantity in realistic cases**. We must always start with either mass or mole balances. These quantities are always conserved (omitting nuclear reactions of course...). What you will find though is that in ideal cases, such as when density is assumed constant, mass balances can break down to volume balances, which makes life simpler. Just be careful that you don’t start with a volume balance.

$$\frac{dM_{total}}{dt} = \sum_i (M_{in,i}) - \sum_i (M_{out,i})$$

Dividing through by the average constant density:

$$\frac{1}{\rho} \cdot \frac{dM_{total}}{dt} = \frac{1}{\rho} \cdot \left(\sum_i (M_{in,i}) - \sum_i (M_{out,i}) \right)$$

$$\frac{dV_R}{dt} = \sum_i (F_{in,i}) - \sum_i (F_{out,i})$$

Naturally, with only one inlet and one outlet, this relation breaks down to our previous assumption of equal flow at steady state.

$$\begin{aligned}\frac{V_R}{dt} &= \sum_i (F_{in,i}) - \sum_i (F_{out,i}) \\ (0) &= F_{in} - F_{out} \\ F &= F_{in} = F_{out}\end{aligned}$$

3. The construction of the mass balances for N and B follow the exact same logic used in question 3. Since we have a constant reactor volume:

$$\begin{aligned}\frac{dM_N}{dt} &= \frac{dNV_R}{dt} = \frac{dN}{dt} V_R = \underbrace{\sum_i (N_{0,i} F_{in,i})}_{\text{inlet}} - \underbrace{\sum_i (N F_{out,i})}_{\text{outlet}} + \text{Generation} - \text{Consumption} \\ \frac{dM_B}{dt} &= \frac{dBV_R}{dt} = \frac{B}{dt} V_R = \underbrace{\sum_i (B_{0,i} F_{in,i})}_{\text{inlet}} - \underbrace{\sum_i (B F_{out,i})}_{\text{outlet}} + \text{Generation} - \text{Consumption}\end{aligned}$$

We have assumed there is no generation of N and no consumption of B (i.e. bacterial death, also known as endogenous decay) in the reactor and we have assumed we have only one inlet and outlet stream and that the inlet stream is sterile. Therefore:

$$\frac{dN}{dt} = N_0 \frac{F}{V_R} - N \frac{F}{V_R} - \left(\frac{1}{Y_B} \right) \left(\mu_{\max} \frac{N}{K + N} \right) B$$

$$\frac{dB}{dt} = -B \frac{F}{V_R} + \left(\mu_{\max} \frac{N}{K + N} \right) B$$

(As a quick little extra bit of information, the factor $\frac{F}{V_R}$ is commonly referred to as the “dilution rate” in bioreactor engineering and is a very important design factor)

4. When considering the steady state of bioreactors it is often best to start with the organism balances (this is because they are more likely to not contain an inlet term and so are simpler to work with). This is also because the organism balances will be your source for determining if the steady state substrate values are directly calculable.

If you cannot explicitly write all of the substrate concentrations from the steady state organism balances then there is no need to derive the rest of the steady state balances in the hopes of an analytical solution. Simply skip

to numerically determining the steady state.

$$\begin{aligned}
 \frac{dB}{dt} &= -B \frac{F}{V_R} + \left(\mu_{\max} \frac{N}{K + N} \right) B \\
 (0) &= -\bar{B} \frac{F}{V_R} + \left(\mu_{\max} \frac{\bar{N}}{K + \bar{N}} \right) \bar{B} \\
 \bar{B} \frac{F}{V_R} &= \left(\mu_{\max} \frac{\bar{N}}{K + \bar{N}} \right) \bar{B} \\
 \frac{F}{V_R} &= \left(\mu_{\max} \frac{\bar{N}}{K + \bar{N}} \right) \\
 \frac{F}{V_R} (K + \bar{N}) &= \mu_{\max} \bar{N} \\
 \left(\mu_{\max} - \frac{F}{V_R} \right) \bar{N} &= \frac{F}{V_R} K \\
 \bar{N} &= \frac{\frac{F}{V_R} K}{\left(\mu_{\max} - \frac{F}{V_R} \right)}
 \end{aligned}$$

$$\begin{aligned}
 \frac{dN}{dt} &= N_0 \frac{F}{V_R} - N \frac{F}{V_R} - \left(\frac{1}{Y_B} \right) \left(\mu_{\max} \frac{N}{K + N} \right) B \\
 (0) &= \frac{F}{V_R} (N_0 - \bar{N}) - \left(\frac{1}{Y_B} \right) \left(\mu_{\max} \frac{\bar{N}}{K + \bar{N}} \right) \bar{B} \\
 \left(\frac{1}{Y_B} \right) \left(\mu_{\max} \frac{\bar{N}}{K + \bar{N}} \right) \bar{B} &= \frac{F}{V_R} (N_0 - \bar{N}) \\
 \bar{B} &= \frac{\frac{F}{V_R} (N_0 - \bar{N})}{\left(\frac{1}{Y_B} \right) \left(\mu_{\max} \frac{\bar{N}}{K + \bar{N}} \right)}
 \end{aligned}$$

5. As was seen in the previous part of the solution, the steady states are indeed directly calculable if given values for F , V_R , N_0 , K , μ_{\max} , and Y_B . It is also interesting to note that the theoretical maximum for the dilution rate is set by the denominator of the steady state version of the organism balance and is equal to the maximum specific growth rate (μ_{\max}). When you think about it this makes perfect sense. If you are putting in flow faster than your organisms can grow then they will simply be washed out.

While less obvious, it can also be shown that \bar{N} will approach the inlet concentration of N_0 as the dilution rate approaches μ_{\max} (This makes intuitive sense because if you are feeding the reactor at a rate that is almost washing out the bacteria then their concentration will be lower and so most of your substrate will exit the reactor). Obviously if your goal is to digest N then you would want to make your reactor as large as possible, but realistically after a certain point there would be diminishing returns for increasing the reactor size. Therefore space, cost, and legal constraints usually set the reactor size.

As an aside, it is interesting to note that if your goal was to grow as much bacteria as possible (this could happen in specialty chemicals for example), and assuming your inlet substrate concentration is much larger than the half saturation constant (I will not go into why this requirement is needed...) then you would actually want to operate the reactor as close as possible to wash out. This is because while the reactor concentration of the bacteria would likely be smaller, the substrate concentration would be much higher and so the overall production rate would be higher.

Question 5 [2]

Note: To review and practice material related to number systems and round-off errors.

1. Convert the following binary numbers to base-10:

- 01010111
- 1110.0001

2. Convert the following hexadecimal numbers to base-10:

- 9A8B7C6D
- DEADB0EF

3. In either MATLAB or Python: write a short function that returns the smallest positive number used on your computer. In other words your computer will be unable to reliably distinguish between zero and a quantity smaller than this number.

4. How does your value (above) compare to what MATLAB or Python reports?

- **MATLAB:**

```
>> help eps % read what this command does
>> eps
```

- **Python:**

```
>>> np.finfo(np.double).eps
```

Solution

1. $(01010111)_2 = 0 \times 2^7 + 1 \times 2^6 + 0 \times 2^5 + 1 \times 2^4 + 0 \times 2^3 + 1 \times 2^2 + 1 \times 2^1 + 1 \times 2^0 = 87$

$$(1110.0001)_2 = 1 \times 2^3 + 1 \times 2^2 + 1 \times 2^1 + 0 \times 2^0 + 0 \times 2^{-1} + 0 \times 2^{-2} + 0 \times 2^{-3} + 1 \times 2^{-4} = 14.0625$$

2. The conversion from hexadecimal to decimal is done similarly to the conversion from binary to decimal. The difference is that the base is 16 here. Also, the letters A to F represent integer numbers 10 to 15 respectively.

$$(9A8B7C6D)_{16} = 9 \times 16^7 + (10) \times 16^6 + 8 \times 16^5 + (11) \times 16^4 + 7 \times 16^3 + (12) \times 16^2 + 6 \times 16^1 + (13) \times 16^0 = 259\,283\,2621$$

$$(DEADB0EF)_{16} = 13 \times 16^7 + 14 \times 16^6 + 10 \times 16^5 + 13 \times 16^4 + 11 \times 16^3 + 0 \times 16^2 + 14 \times 16^1 + 15 \times 16^0 = 3\,735\,924\,975$$

3. The following code obtains the smallest positive number.

MATLAB code

```
%This code returns the smallest positive number which can be obtained as
%the distance from 0.0 to the next larger double precision number.
```

```
number = 1.3; % initialize to any positive value as long as
              % it is not already below the under flow limit
while (number>0)
    epsilon = number; % store current value of number
    number=number/2; % keep dividing by two
end
```

```
fprintf('\n My epsilon =%d\n',epsilon) %Solution: epsilon = 4.940656e-324
```

```
%Compare your epsilon with the one returned by the MATLAB function >>eps(0)
fprintf('\n Epsilon by MATLAB =%d\n',eps(0))
```

Python code

```
# This code returns the smallest positive number which can be obtained as
# the distance from 0.0 to the next larger double precision number.

number = 1.3                                # initialize to any positive value as long as
                                           # it is not already below the under flow limit

while (number>0):
    epsilon = number                         # store current value of number
    number = number/2.0                     # keep dividing by two

print('\n My epsilon = %g \n' % epsilon) # epsilon = 4.94066e-324

# Python does not have an eps(0.0) function
# So Python users can ignore this part of the question.
```

Recall that most computers internally convert base 10 input values to binary, base 2 values before performing any calculations. Therefore, the initial value of *number* in the above code is approximated to the nearest binary representation in the computer. Then the division of that binary representation by two will always have an exact binary representation. Therefore, unlike the initial value of *number* which may not be finitely representable in binary, its subsequent values obtained through the while loop are divisible by two and thus exactly representable in binary. That is why, the variable “epsilon” must eventually reach the smallest positive number $.0000\dots0001 \times 2^{-111\dots111}$.

The code keeps dividing *number* by two in order to make it closer and closer to zero. At some point, the current value of *number* cannot be distinguished from zero (the condition in the while loop). This is where the algorithm stops. We report the value right before the current one as the smallest positive number that can be distinguished from zero. This previous value is always stored as *epsilon* in the above code.

The epsilon value around 0.0 returned by MATLAB matches that computed by our code. Notice that this value is not explicitly classified as the smallest *normalized* representation (in computers, binary numbers are normalized to have a lead digit of 1). This is because some computers allow for “gradual underflow”. This means that some computers allow the mantissa to become denormalized (i.e. have 0 as a lead digit) once the exponent has reached its smallest allowable value (thereby making the gap between 0 and the epsilon smaller). As an aside, note that your solution will likely equal `realmin()*eps()`. There is a very logical reason for this (left to you as an exercise to see why!).

END