Stochastic Dynamics: Random Walks, Recurrence Relations, Stochastic Maps/Evolutions [1]

Scott A. Strong

To portray your life in order would be absurd: I remember you at random. My brain resurrects you through stochastic details, like picking marbles out of a bag.
[Édouard Levé (1965-2007)]

Pseudorandom number generation is a deterministic process that yields sets of numbers that behave as random numbers.[2] True random number generation leverages physical processes that are necessarily random to us. For example, in the movie Contact, Jodie Foster portrays the fictional astronomy, Dr. Eleanor "Ellie" Arroway, listening to radio signals detected by the Karl G. Jansky Very Large Array (VLA) located in central New Mexico on the Plains of San Agustin, which is a centimeter-wavelength radio astronomy observatory associated with many key scientific discoveries, e.g., ice on Mercury, observations of supermassive black holes, microquasars, Einstein rings, and gamma-ray bursts. The movie, based on the 1985 Carl Sagan book of the same name, is about the discovery of extraterrestrial life and our first contact. Early in the movie, we find Ellie listening to the audio representation of signals detected by the VLA. Hearing a structured pattern in the signal causes her to jump into her light blue '68 Chevy Impala and race towards the observatory offices. Theatrics aside, the radio emissions from her spark plugs firing would have significantly compromised the signal detected by the VLA. Perhaps, with the knowledge of the periodic firing, this could have been accounted for and filtered out. However, considering the 6.8 million lightning strikes per day, causing electromagnetic emission (noise), ultimately manifesting from fluid motion constrained to a rotating ball, we have to honor the fact that even if a process is deterministic, that doesn't imply we possess non-trivial predictive powers from its associated model. In addition, some amount of existing background noise can be attributed to the primordial universe.[3] Fundamentally, unwanted noise is a natural part of observing deterministic processes, and the further we are from a laboratory setting, the more we encounter it.

Even within laboratory settings, random fluctuations may be intrinsically tied to the observed phenomenon. In 1827 the botanist Robert Brown (1773-1858) observed irregular motion of particles, amyloplasts (starch organelles) and spherosomes (lipid organelles), ejected from the grains of pollen from the Clarkia Puchella while suspended in hydrogen-hydroxide.[4] Furthermore, he observed this in other suspended inorganic matter, and while it led him to conclude that the particles were not *alive*, it did not lead to a theoretic explanation of the observed motion. Later, in the doctoral thesis of Louis Jean-Baptiste Alphonse Bachelier, a student of Henri Poincaré, defined the first mathematical model of Brownian motion (1900) and applied it to value stock options. While setbacks due to World War

[1] Last compiled: Monday 21st November, 2022 at 23:21:26.

[2] The exact details of how this is accomplished are non-trivial and worth a bit of discussion. First, we would like there to be a high probability that elements from the sequence are different from one another, that the sequence of numbers behave similarly to random sequences under specific statistical tests, that one cannot get the sequence from subsequences, and that knowledge of the sequence does not lead to knowledge of the initial value. One of the first procedures was presented by John von Neumann in 1946 and is known as the middle square algorithm, which starts with a seed(initial) value that is then squared prior to having its middle four digits taken to be the next random number in the sequence. This process will eventually cycle, and so a chief concern is efficiently extending cycle length.

[3] "Tune your television to any channel it doesn't receive and about 1 percent of the dancing static you see is accounted for by this ancient remnant of the Big Bang. The next time you complain that there is nothing on, remember that you can always watch the birth of the universe." Bill Bryson, A Short History of Nearly Everything (2003)

[4] This plant was harvested in the Pacific Northwest east of the Cascades, and is in the family of primrose, which also contains fuchsia and fireweed.

I, and academic politics, kept Bachelier from developing his theories further, Albert Einstein would rederive and extend them in 1905, leading to his first major scientific contribution connecting the kinetic theory of heat with the motion of small particles suspended in a stationary liquid, providing an explanation for the empirical evidence of atomic theory.[5] Later, in 1926, Jean Baptiste Perrin would conduct experiments verifying Einstein's theory, which would also determine Avagadro's number asserting a discontinuous structure to matter, leading to a Nobel Prize. After another roughly twenty-year period, the mathematical theory of stochastic differential equations was invented by Kiyosi Itô and allowed the study to progress past linear equations. [6]

[5] For a bit of historical context, on 12/3/2021, Kane Tanaka was the oldest human on planet Earth, born in 1903. She would have received an elementary school education as post-war Japan had taken a keen interest in societal rebuilding. Likely, the curriculum would have emphasized reading, writing, and arithmetic but may have included history, geography, and natural science. Before World War II, secondary education was male only in Japan, so questions related to molecules would only have occurred in adult life. The term molecule, popularized by Avagadro in 1811, might have appeared in secondary education, but at this point in America, 18% of $15-18$-year-olds were enrolled, and 9% of 18-year-old citizens would attain this degree.

[6] In 1956, AP biology, chemistry, and physics were introduced in the United States as we ramped up science efforts during the space race. This means that my father, who would have been in middle school around this time, may have heard the word molecule thrown around. My mother, who was raised in Taiwan under Chinese Nationalist rule, would have only been mandated six years of primary education. As with many other women in her position, she received less than this.

1 *Random Walks on Lattice and Graphs*

> A drunk man will find his way home, but a drunk bird may get lost forever.
>
> Shizuo Kakutani (1911-2004)

A random walk is a stochastic process defining a time series corresponding to a path through a mathematical space. If we let $X_1, X_2, X_3, \ldots, X_n$ be identically distributed random variables in $\mathbb{R}^n$, sampled from a population with a well-defined mean, μ_X, and variance, σ_X^2, then the sequence of positions $\{S_i\}_{i=1}^n$ is given by,

$$\begin{aligned} S_n &= S_{n-1} + X_n && (1)\\ &= S_{n-2} + X_{n-1} + X_n && (2)\\ &= S_{n-3} + X_{n-2} + X_{n-1} + X_n && (3)\\ &= S_0 + X_1 + X_2 + X_3 + \ldots X_n && (4)\\ &= S_0 + \sum_{i=1}^{n} X_i, && (5) \end{aligned}$$

which tells us that at the $n^{th}-$step, the walker is located a distance away from the starting location S_0 defined by the sum of the random steps. The backbone of traditional statistical theory, known as the central limit theorem, tells us that for large n, the paths are asymptotically normal, with mean $n\mu_X$ and variance $n\sigma_X^2$. Consequently, samples of long paths can be statistically differentiated through hypothesis testing.

Some of the contexts in which random walks appear. 1) evolution of stock prices, 2) genetic drift where gene variants arise/disappear randomly, 3) sequences of randomly oriented rigid monomer rods leading to amino acids in a polypeptide, 4) movement in motile bacteria, 5) size estimation in large networks like the internet.

Before we study random walks, we present a heuristic argument for the quote leading this section. First, a random walk is called *recurrent* if it visits its starting position infinitely often with probability one, and *transient* if it is not recurrent. Using this language, we rephrase the quote as:

- Random walks in $\mathbb{R}^n$ are transient for $n \geq 3$.

The rationale for this relies on the convergence to the normal distribution given by the central limit theorem.

Suppose that you are walking in one dimension on the integers with steps defined by i.i.d. random variables sampled from a population with well-defined mean and variance. Assuming $n \gg 1$, at the $n^{th}-$step, the bulk of walkers will be at $n\mu_X$ with mean squared distance away from this location given by, $\sqrt{n}\sigma_X$. The probability of a given walker being on the interval $[n\mu_X - 3\sigma_X\sqrt{n}, n\mu_X + 3\sigma_X\sqrt{n}]$ is about 99.7%. For ease of use, let $\mu_X = 0$ and $\sigma_X = 1$ and ask, "what is the probability that the walker is at the origin for a given step when $n \gg 1$?" There are about $2 \cdot 3\sqrt{n}$ lattice points on this interval, and if we assume the probability of being at any one of them is uniform, then the probability that the walker is at the origin scales like $1/\sqrt{n}$. Next, we define the indicator function $\mathbf{1}_{S_n=0}$, which returns 1 if the walker is at the origin and zero if it is not, and we would like to determine the expected number of times the walker visits the origin. Since the expectation is linear in its argument and the expectation of the indicator function is just the probability of the event, we have,

$$E\left[\sum_{n=1}^{\infty} \mathbf{1}_{S_n=0}\right] = \sum_{n=1}^{\infty} E\left[\mathbf{1}_{S_n=0}\right] \tag{6}$$
$$= \sum_{n=1}^{\infty} \mathbb{P}(S_n = 0) \tag{7}$$
$$\propto \sum_{n=1}^{\infty} \frac{1}{\sqrt{n}}, \tag{8}$$

which is a p-series with $p = 1/2$. Thus, the divergence indicates that we expect to return to the origin an infinite number of times, i.e., a random walk on $\mathbb{Z}$ is recurrent. Under the same assumptions, a two-dimensional walker has a 1 in $(\sqrt{n})^2$ chance of being at the origin for a large but arbitrary step. For walks on $\mathbb{Z}^d$ we have that the expected number of times the walker returns to the origin is,

$$E\left[\sum_{n=1}^{\infty} \mathbf{1}_{S_n=0}\right] = \sum_{n=1}^{\infty} E\left[\mathbf{1}_{S_n=0}\right] \tag{9}$$
$$= \sum_{n=1}^{\infty} \mathbb{P}(S_n = 0) \tag{10}$$
$$\propto \sum_{n=1}^{\infty} \frac{1}{\sqrt{n}^d} \tag{11}$$
$$= \sum_{n=1}^{\infty} \frac{1}{n^{d/2}}. \tag{12}$$

Since the p-series only diverges for $d \leq 2$, walks in three dimensions and higher cannot be recurrent and drunk birds may have to relocate.

1.1 *Simulation and analysis of random walks on $\mathbb{Z}$: Gambler's Ruin*

> Luck is a very thin wire between survival and disaster, and not many people can keep their balance on it.
>
> Hunter S. Thompson (1937-2005)

DUTCH MATHEMATICIAN, PHYSICIST, ENGINEER, ASTRONOMER, AND INVENTOR, Christiaan Huygens (1629-1695) became interested in games of chance after he visited Paris in 1655 and encountered the work of Fermat, Blaise Pascal and Girard Desargues. His work in this area would define expectation value and directly influence Jacob Bernoulli's work on probability theory. In what is regarded as the first published work on probability theory, *On Reasons in the Game of Chance* (1657), Huygens discusses several problems, and in what is known as his fifth problem, he considered a two-player game of chance which terminates when one of the two players runs out of the resources needed to initiate the game through a bet. This gedankenexperiment is now known as the gambler's ruin and can be modeled as a random walk on the one-dimensional lattice, $\mathbb{Z}^1$.

The simplest possible case is a single event/choice game at a casino backed by an infinite amount of money. An example of this could be betting on a roulette wheel, where each bet costs \$1 and the player starts off with \$$n$ and has a probability, $p \in (0,1)$, of winning against the unbreakable bankrolled house. This is nothing more than our Rademacher random variable acting on the sample space of win/loss to produce the events $\$1/-\$1$, with probabilities defined by the discrete density $\rho(X=\$1) = p$ and $\rho(X=-\$1) = 1-p$. Moreover, this process is a simple random walk on $\mathbb{Z}$ but asymmetric because, at the end of the day, it is a big hustle favoring the house.

For the asymmetric case, if we define the winning pot, \$$w$, as the initial bankroll, \$$n$, plus the desired increase in money, \$$m$, then it can be shown that the probability a gambler reaches \$$w$ before going bust is given by,

$$P_n = \frac{\left(\frac{1-p}{p}\right)^n - 1}{\left(\frac{1-p}{p}\right)^w - 1}. \tag{13}$$

If we start with a large bankroll, i.e., $n \gg 1$, then we have the approximation,

$$P_n \approx \frac{\left(\frac{1-p}{p}\right)^n}{\left(\frac{1-p}{p}\right)^w} = \left(\frac{1-p}{p}\right)^{n-w} = \left(\frac{1-p}{p}\right)^{-m}. \tag{14}$$

If the gambler is playing American roulette, i.e., 18 numbered red/black outcomes with additional green 00 and 0 outcomes, by consistently betting on black, then $p = 18/38$ and the odds of coming out \$100 ahead, calculated from $(10/9)^{-m}$, is about 1 in $37,649$. So, even if you start with exceptionally large stacks, you are unlikely to come out with an extra Benjamin.

We might also want to know if we should expect to even see this outcome before going bust. For this, we need to calculate the first passage, *hitting time*, or *stopping time* for the discrete-time stochastic process. It can be shown that if H_n is the expected number of bets the gambler can make before going home broke or a winner, then

$$H_n = \frac{n}{1-2p} - \frac{w}{1-2p}\left[\frac{\left(\frac{1-p}{p}\right)^n - 1}{\left(\frac{1-p}{p}\right)^w - 1}\right] \tag{15}$$

$$= \frac{n}{1-2p} - \frac{w}{1-2p}P_n, \tag{16}$$

where $H_0 = 0$ and $H_w = 0$. If the house has an advantage and the amount we hope to gain is large, then $P_n \to 0$ as $m \to \infty$ and we can approximate the hitting time as

$$H_n \approx \frac{n}{1-2p}. \tag{17}$$

While we can derive these formulae by building the associated recurrence relations, we begin by simulating this process to see the associated results manifest from the statistics of many gamblers. For one gambler we have the following,

```
for(i in 1:numberOfGamblers) #outer loop iterating over each gambler
  {
    # Initialization of internal loop variables
    betCount = 0;
    winner   = 0;
    loser    = 0;
    n = initialBankRoll; # initialBankRoll defined outside of loops

    while(betCount < maxNumberOfBets) # Betting loop(capped) for each gambler
    {
      betCount = betCount + 1;
      nNew = n + sample(c(-1,1), size = T, replace = TRUE, prob = c(1-p, p));
      n = nNew;

      if(n == 0) # Control statement to record data if gambler goes bust
      {
        loser=1;
        timeToLose[i] = betCount;
        break; # Gambler went bust, exit to outer loop
      }

      if(n == (initialBankRoll+desiredWinnings) )# record if gambler wins
      {
        winner=1;
        timeToWin[i] = betCount;
        break; # Gambler wins, exit to outer loop
      }
    } #End inner loop simulating gambler's run
    potAtEnd[i] = n; # Records winnings of each gambler.
  } #End outer loop iterating over gamblers
```

Prior to this, we define the variables `initialBankRoll`, `desiredWinnings`, `numberOfGamblers`, `maxNumberOfBets`, `potAtEnd` in addition, two vectors containing relevant results to the group of gamblers, i.e., `timeToLose` and `timeToWin`. After this, we calculate and print the descriptive statistics, `meanTimeToBust`, `percentageOfBusts`, `meanTimeToWin`, and `percentageOfWinners`, for review.

To calculate the statistics necessary to verify the probability of going bust as well as stopping times, we nest the previous code in an outer loop that scopes over a specified `numberOfseeds`, which are randomly selected.

```
for (j in 1:numberOfseeds) { #Outer loop runs several groups of gamblers
  set.seed(seeds[j]) #Update the seed
  # Re-initilization of vectors to hold win/loss/pot data
  timeToLose = matrix(,numberOfGamblers);
  timeToWin  = matrix(,numberOfGamblers);
  potAtEnd   = matrix(,numberOfGamblers);

  # Gambler simulation code goes here.
  #  code calculates timeToLose, timeToWin vectors for a given simulation

  distributionOfWins[j] = 100*sum(!is.na(timeToWin))/numberOfGamblers;
  distributionOfTimeToLose[j] = mean(timeToLose, na.rm = TRUE);
}
```

The problems below will use several variants of this code base to realize these properties of the gambler's ruin.

Problems:

1. In the shared drive, you will find `GamblersRuin_Problem1.R`, which simulates repeated black/red betting for American roulette for an ensemble of gamblers who are hard-limited to a specific number of bets regardless of their current holdings. It calculates and stores the `timeToLose`, `timeToWin`, `potAtEnd` for each of the gamblers. Run this code and replicate the test cases given in its comments.

2. Assume that the gambler repeatedly plays red on an American roulette table and starts with a large amount of money, i.e., $n \gg 1$. How much money should they look to increase their wealth by having a 1% chance of making this happen?

3. Confirm the results of the previous problem with the R code called `GamblersRuin_Problem3.R`. Start by providing the gambler an `initialBankRoll`, `desiredWinnings`, and probability of winning to 100 gamblers who are able to only make 5000 bets. Run the code and summarize the statistics of `distributionOfWins` and plot a histogram for `distributionOfWins`.

4. A roulette round can take somewhere between three and six minutes. Suppose that a gambler wishes to make one hundred billion dollars and will leave only after they do this or go broke. If they only have three hours to spend and plan to bet $1 per spin, assuming the roulette spins take 4 minutes, then how much money should they arrive with?

5. Confirm the results of the previous problem with the R code called `GamblersRuin_Problem5.R`. Start by providing an `initialBankRoll`, `desiredWinnings`, and

probability of winning to 100 gamblers who are able to only make 5000 bets. Run the code and summarize the statistics of `distributionOfTimeToLose` and plot a histogram for `distributionOfTimeToLose`.

1.2 *Simulation and analysis of random walks on $\mathbb{R}^3$: Photon's escape from the Sun*

> We exist in a bizarre combination of Stone Age emotions, medieval beliefs, and god-like technology.
>
> Edward Wilson (1929-2021)

HYDROGEN FUSION TAKES IN FOUR HYDROGEN NUCLEI (PROTONS) AND TWO ELECTRONS TO GENERATE A HELIUM NUCLEUS, TWO NEUTRINOS, AND SIX PHOTONS. These photons are then left to bounce around interacting[7] with other massive charged particles along the way. This photonic pinball can be modeled as a random walk in $\mathbb{R}^3$, we can ask the question of how many steps it takes to exit the star so that it can make the eight-minute journey to Earth. In photonic pinball, one goes to great lengths to define the *mean free path*, which is the length the photon will go before it collides with the charged matter within the plasma and is often larger than the average distance between particles. For example, in our Sun, the average density of matter is 1.408 grams per cubic centimeter, with each parcel holding Avogadro's number of hydrogen atoms. If these atoms were arranged on a lattice, then we would need in-line lattice spacing of 1.056539×10^{-8} centimeters. However, accounting for electromagnetic effects, a photon tends to only collide with matter every centimeter.[8] This is our assumed mean free path of the photon.

[7] The photon, a quanta of the electromagnetic field, interacts with the proton, a charged particle capable of interacting with the electromagnetic field through absorption and subsequent re-emission. The photon leaving is technically not the same photon that was absorbed, but this does not undermine our random walk hypothesis.

[8] There are actually a variety of results for this parameter. I've chosen this since it sets a scale so that each step is one centimeter.

We model these successive collisions as a random walk through $\mathbb{R}^3$ with new spherical variables $\theta \in [0, \pi]$ and $\phi \in [0, 2\pi)$ chosen after each interaction/step, with radial displacement 1 centimeter between hits. We will use the code `PhotonRW.R` whose base is given by the following.

```
while(steps<stepTotal) #Loop to simulate RW up to a maxSteps
{
  randomTheta=runif(1, min=0,max=pi); #generate a random polar angle
  randomPhi = runif(1, min=0,max=2*pi); #generate random azimuthal angle
  # define (x,y,z) motions by standard formulae
  x = meanFreePath * sin(randomTheta)*cos(randomPhi);
  y = meanFreePath * sin(randomTheta)*sin(randomPhi);
  z = meanFreePath * cos(randomTheta);
  SNew = Sn + c(x,y,z); #update the three-space location of the walker
  Sn=SNew;
  steps = steps +1;
  radialDistance = sqrt((Sn[1])^2+(Sn[2])^2+(Sn[3])^2); #compute the radial dist.
}
```

While this is a good approach, conceptually, it will take far too long to complete, even for a straight-line radial escape of the photon from the Sun's core. The Sun's radius is 6.957×10^{10} centimeters and if one iteration of this loop takes 4.196167×10^{-5} seconds, then a single photon moving radially outward would need about a month of compute time to make it out and we want to look at the statistics of ensembles of photon walkers! Adding to this is the inefficient exploration associated with random walking and we have a significant problem.

Let $\underline{R}_n$ be the random variable defining the radial distance away from the center of the Sun's core, taken to be the origin of $\mathbb{R}^3$, which is defined by the walk,

$$\underline{R}_{n+1} = \underline{R}_n + \underline{X}_n(\theta_n, \phi_n), \quad n = 1, 2, 3, \ldots, \tag{18}$$

where where θ_n and ϕ_n are independent random variables sampled uniformly over $[0, 2\pi]$, and $[0, \pi)$, respectively. The central limit theorem tells us that $|\underline{R}_n| \sim \mathcal{N}(n\mu_X, n\sigma_X^2)$ so that at the $n^{th}-$step we expect that the majority of photons are within $\sqrt{n}$ centimeters from the center of the Sun, let d be this distance. From this, we can estimate the number of steps before this distance equals the radius of the Sun, i.e., $d = 6.957 \times 10^{10} = \sqrt{n}$, and the edge of the majority would require roughly 10^{20} steps for escape or about 6.5 billion years of compute time.[9]

[9] For a frame of reference here, the universe is approximately 14 billion years old and our solar system has only been around for a third of this time.

Problems:

1. Using `PhotonRW.R` found in the shared drive, send 100 photons out from the center of the Sun to collide with the charged matter with a mean free path of 1cm, but limit the simulation to 10,000 steps. Looking at the histogram, what do you notice? How does this result relate to the central limit theorem? Run `summary(distanceFromOrigin)` to quantify your suspicions with proper numbers.

2. If the bulk of our photons made it 100 cm during our 10,000 steps, then estimate how many more steps it will take to get to the edge of the Sun, Specifically, calculate `timePerStep*stepTotal*sunRadius/distanceFromOrigin`. Is this a long time?

3. Watch Science Shorts - How Long Does It Take for a Photon to Leave the Sun? (`https://youtu.be/79SG_2XHl_I`). How long do they say it takes a photon to travel from the core to the surface? Note, it is interesting that after the radiative zone, the decreasing density of the Sun allows convection to dominate over diffusion more efficiently transporting the photons radially outward. We do not take this into account, but still, our difference in time scales holds.

4. At this point, you should have a predicted value for the time to escape the Sun which is much lower than what the experts have

said. So, what has happened? Recall that the central limit theorem tells us that the root mean squared distance is proportional to $\sqrt{n}$, i.e., $R_{rms} = \sqrt{n}$. If we consider n to be a continuous variable, then we can define its velocity as $\dot{R}_{rms} \propto 1/\sqrt{n}$, which decays in time. With this in mind, why was our estimate for escape time low compared to what the experts have said? Transform the time prediction to the number number of steps. How does this compare to the prediction given in above?

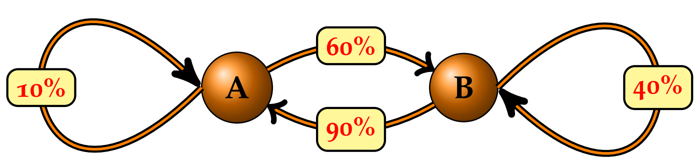
5. There exist several different estimates for the mean free path of a photon. It turns out that this isn't a quantity that astrophysicists care about because it doesn't relate to the relevant studied phenomenon. Suppose that the mean free path was on the order of 10^{-3} centimeters. How does this affect the predicted time of photon escape?

1.3 *Discrete-Time Markov chains*

> Don't think of the overwhelming majority of the impossible.
>
> Ramtin Alami's Markov chain (1999-present) [See *How I generated inspirational quotes with less than 20 lines of python code*, published on Medium (2018)]

THE RANDOM WALK, AS WE HAVE DEFINED IT, IS AN EXAMPLE OF A DISCRETE-TIME MARKOV CHAIN, i.e., a memoryless stochastic process[10] and in this section, we will explore them by conducting random walks along graphs. For example, we could have the following graph representing the percentage of walkers positioned at either state A or state B, that choose to go to state A or state B at the next time step.

[10] Assuming a discrete state space with the power-set sigma algebra indexed by the natural numbers, memoryless is formulated as $P(X_n = x_n | X_{n-1} = x_{n-1}, \dots, X_0 = x_0) = P(X_n = x_n | X_{n-1} = x_{n-1})$, which is saying that the probability that the random variable takes on the state x_n at the n^{th}-step, given all of the preceding outcomes, is equivalent to the probability of the event x_n given only the last outcome. Another way to phrase this is that the time we must wait until a certain event does not depend on how much time has elapsed.



We use this to define the following two-by-two matrix representing the transition probabilities defined by this graph,

$$\mathbf{P} = \begin{bmatrix} 0.1 & 0.6 \\ 0.9 & 0.4 \end{bmatrix}, \tag{19}$$

where the first column/row represents A and the second column/row represents B. The second row, first column, element defines the probability that a walker on the graph in state B transitions to state A.

Notice this square matrix has columns that sum to one which implies that it is a *stochastic matrix*. Also, for some positive power, e.g., $\mathbf{P}^1$, all of the entries are non-zero implying that the stochastic matrix is also *regular*. It is known that if the transition matrix for a Markov chain is regular, then every initial distribution of walkers on the graph tends to a unique limit distribution known as the steady-state vector.

In terms of mathematics, we have if $\mathbf{x}_0 \in \mathbb{R}^2$ is a stochastic vector, then application of the transition matrix, $\mathbf{P}$, results in a new discrete-time updated vector consistent with the transition probabilities, i.e., $\mathbf{P}\mathbf{x}_0 = \mathbf{x}_1$. If the steady-state vector, $\mathbf{q} \in \mathbb{R}^2$, is a fixed-point of the map, i.e., $\mathbf{P}\mathbf{q} = \mathbf{q}$, then it is an eigenvector of $\mathbf{P}$ associated with eigenvalue $\lambda = 1$. In our case,

$$(\mathbf{P} - \lambda\mathbf{I})\,\mathbf{q} = \begin{bmatrix} -0.9 & 0.6 \\ 0.9 & -0.6 \end{bmatrix} \begin{bmatrix} q_1 \\ q_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \implies \mathbf{q} = \begin{bmatrix} 2/5 \\ 3/5 \end{bmatrix}, \tag{20}$$

Theory tells us that **every** initial distribution tends to this steady-state vector as the number of iterations goes to infinity. While we have the steady-state vector, we haven't yet shown that continued reapplication of the transition matrix will result in this state.

To see that every initial distribution tends to this steady-state vector, we recognize that if the transition matrix has a complete set of eigenvectors, then it can be made diagonal relative to its eigenbasis, i.e., $\mathbf{P} = \mathbf{U}\mathbf{D}\mathbf{U}^{-1}$, where the coordinate change matrix $\mathbf{U}$ is populated with the two eigenvectors of $\mathbf{P}$ and $\mathbf{D}$ is a diagonal matrix whose columns are the corresponding eigenvalue multiplies of the standard basis vectors. This is a fundamental result in linear algebra. It says that when a diagonalization is available, there exists a basis in which the linear system of equations appears completely decoupled. That is, row-reduction is unnecessary since the diagonal matrix is already in an echelon form. Or, in terms of transformations of the plane, there is a specific direction in $\mathbb{R}^2$ that is neither rotated nor scaled by the linear transformation represented by $\mathbf{P}$. The positive and normalized vector on this untransformed line in $\mathbb{R}^2$ is our steady-state vector. For our needs, we must notice that since diagonal matrices act as much like scalars as one could hope, the diagonal decomposition of $\mathbf{P}$ gives us a means to represent the powers of the transition matrix,

$$\mathbf{x}_n = \mathbf{P}^n\mathbf{x}_0 = (\mathbf{U}\mathbf{D}\mathbf{U}^{-1})^n\mathbf{x}_0 = \mathbf{U}\mathbf{D}^n\mathbf{U}^{-1}\mathbf{x}_0, \tag{21}$$

where powers of the diagonal matrix are calculated by raising the diagonal elements to the associated power. In our case, we have the explicit calculation,

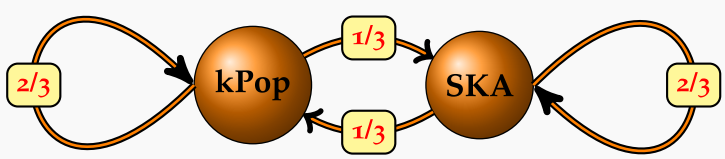
$$
\begin{aligned}
\lim_{n\to\infty} \mathbf{U}\mathbf{D}^n\mathbf{U}^{-1}\mathbf{x}_0 &= \begin{bmatrix} 2/5 & 1 \\ 3/5 & -1 \end{bmatrix} \begin{bmatrix} \lim\limits_{n\to\infty} 1^n & 0 \\ 0 & \lim\limits_{n\to\infty} (-.5)^n \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 3/5 & -2/5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \\
&= \begin{bmatrix} .4 & .4 \\ .6 & .6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} .4(x_1+x_2) \\ .6(x_1+x_2) \end{bmatrix} \\
&= \begin{bmatrix} .4 \\ .6 \end{bmatrix} = \begin{bmatrix} 2/5 \\ 3/5 \end{bmatrix} = \mathbf{q}.
\end{aligned}
$$

Through the diagonalization, we were able to actually take the limit of infinite reapplication of the transition matrix to an arbitrary initial stochastic vector and find the result is always the steady-state vector. In the context of our weighted graph, regardless of the initial distribution of walkers, in the long-time limit, 40% of the walkers will be found on node A and 60% on node B.

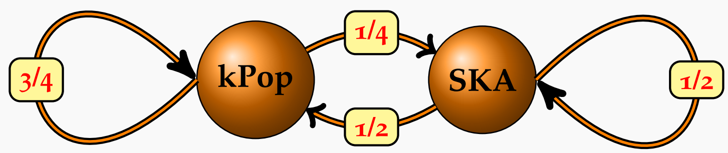
Problems:

1. Watch the first seven minutes of the video Can a Chess Piece Explain Markov Chains? | Infinite Series (`https://youtu.be/63HHmjlh794`) and using the linear algebra above, justify their results for the stationary distributions for the following graphs.

Case 1: Stationary Distribution, kPop= 1/2, SKA=1/2



Case 2: Stationary Distribution, kPop= 2/3, SKA=1/3



2. ✍ For both graphs, using matrix diagonalization, show that the steady-state distribution is the result of infinite reapplication of the transition matrix onto the initial state vector. **Note**: You don't have to do this. You can just trust that it does. At the very least, you should make sure you can read the argument given above.

2 *Recurrence Relations*

> I actually keep having this one recurring dream where I'm a little number standing in a line of other numbers that look identical to me. Then there are more and more of these numbers that follow me, again and again and again. It's more of a nightmare.
>
> Taika Waititi (1975-present)

WITHOUT CALCULUS, WE WOULD NOT HAVE A FRAMEWORK TO UNDERSTAND THE IMPLICATIONS OF NEWTIONIAN MECHANICS AND ELECTRODYNAMICS, as they were the motivation and application of the field. David Hilbert (1862-1943) sought to develop connections between the analysis of differential and integral operators acting on spaces of functions and the characteristic decomposition of linear transformations of finite-dimensional spaces. His infinite-dimensional generalization of Euclidean space led to significant advances in our analysis of functionals and quantum mechanics in the early $20^{th}-$century.

That said, with calculus comes non-trivial heartache as we wrestle with the formalized concepts of limits and convergence, which arise when looking at *instantaneous* change or *continuous* summation. Luckily, we can often take steps to avoid limiting in the continuum, e.g., considering the time between heartbeats instead of the corresponding waveform displayed on an electrocardiogram. For such a pattern, we might consider mapping current states to future stats with a *recur-*

rence relation,

$$y_{n+1} = f(n, y_n), \;\; n = 1, 2, 3, \ldots, \;\; y_0 \in \mathbb{R}, \tag{22}$$

where f is the map that assigns a value to the $(n+1)-$element of the sequence of states by applying the rule/map to the previous state, y_n, and time n. This recurrence relation is first order. When the map depends on two previous values, then we say the recurrence relation is second-order. We say the map is autonomous if there is no explicit reference to n. If the map is linear in the state variables, then it will map linear combinations to linear combinations, and we say the recurrence relation is linear. As you can see, there is a lot of shared language between this discrete-time deterministic map and continuous-time dynamics arising from ordinary differential equations. What is missing, however, is the overt demand to use calculus.

As we have seen before, the linear theory is particularly tractable. For example, given the first-order linear recurrence relation is given by $y_{n+1} = a_n y_n$. If the initial state, y_0, is known, then continued application of the map yields,

$$\begin{aligned}
y_{n+1} &= a_n y_n && (23)\\
&= a_n(a_{n-1} y_n) && (24)\\
&= a_n a_{n-1}(a_{n-2} y_{n-2}) && (25)\\
&= a_n a_{n-1} a_{n-2} \ldots a_0 y_0 && (26)\\
&= y_0 \prod_{i=0}^{n} a_i. && (27)
\end{aligned}$$

Assuming this first-order recurrence relation is also autonomous gives the simplification of $y_n = y_0 a^n$. If the autonomous problem is also inhomogeneous, then we find

$$\begin{aligned}
y_{n+1} &= a y_n + b && (28)\\
&= a(a y_{n-1} + b) + b && (29)\\
&= a(a(a y_{n-2} + b) + b) + b && (30)\\
&= a^{n+1} y_0 + b \sum_{i=0}^{n} a^i && (31)\\
&= a^{n+1}\left(y_0 - \frac{b}{1-a}\right) + \frac{b}{1-a}, && (32)
\end{aligned}$$

which is the sum of a homogeneous solution with a particular solution, assuming $a \neq 1$. This is not unexpected, and if $a = 1$, then we find $y_n = y_0 + bn$.

The analysis for nonlinear maps, or non-autonomous linear maps, is not nearly as straightforward, though it's easy to compute with a simple looping procedure. We will consider nonlinear problems after the following section, where we state some facts about second-order linear constant coefficient recurrence relation, which we use to address some open questions about the time series defined by our gambler.

2.1 *A brief overview of second-order linear autonomous recurrence relations and the gambler's ruin*

> Like Hooke, Huygens made fundamental improvements to the clock as a time-keeping mechanism; and Hooke invented the first passable escapement for the same purpose. ...Huygens discovered the rings of Saturn, and the formula for centrifugal force. He did important work in mechanics and optics, and one of his merits was that he made young Leibnitz enthusiastic for these subjects.
>
> Jacob Bronowski (1908-1974) The Common Sense of Science (1951) "The Scientific Revolution and the Machine."

THE LINEAR THEORY OF RECURRENCE RELATIONS COMPARES QUITE NICELY TO LINEAR ORDINARY DIFFERENTIAL EQUATIONS. Specifically, if the homogeneous equation has constant coefficients, then the general solution is the linear combination of two linearly independent solutions, and these solutions can be found by guessing an exponential function of an unnatural base. For example, if we assume $y_n = \lambda^n$, then

$$ay_{n+2} + by_{n+1} + cy_n = \lambda^n \left(a\lambda^2 + b\lambda + c\right) = 0, \tag{33}$$

where $2\lambda_\pm = -b \pm \sqrt{b^2 - 4ac}$ and

$$y_n = c_1\lambda_+^n + c_2\lambda_-^n, \tag{34}$$

when $\lambda_+ \neq \lambda_-$, which is stable if $|\lambda_\pm| < 1$.[11] As was the case in ordinary differential equations, we are left to sort out what happens when $\lambda_\pm \notin \mathbb{R}$ and $\lambda_+ = \lambda_-$. If $\lambda_\pm \in \mathbb{C}$, proper, then $\lambda_\pm = \alpha \pm \beta i = re^{\pm i\theta}$ where $r^2 = \alpha^2 + \beta^2$ and $\tan(\theta) = \beta/\alpha$. Consequently, we have

$$y_n = k_1 r^n e^{in\theta} + k_2 r^n e^{-in\theta}, \tag{35}$$

and we have stability if $r = |\lambda_\pm| < 1$ with pure oscillations only if $r = 1$. Notice that if $\lambda_\pm = \alpha \pm \beta i$, then $|\lambda| = r = \sqrt{\alpha^2 + \beta^2}$. Additionally, if $\lambda_+ = \lambda_-$, then it is possible to show that $y_n = n\lambda^n$ is a second linearly independent solution, and we have the general solution $y_n = c_1\lambda^n + c_2 n\lambda^n$, in the repeated root case.

So, what does all of this say? Well, if we think about our equation as, $y_{n+2} = f(y_n, y_{n+1}) = -by_{n+1} - cy_n$, then we can identify $(y_n, y_{n+1}) = (0,0)$ as a fixed point of the second-order map and, based on the above analysis, we conclude that if a map has an eigenvalue such that, $|\lambda| > 1$, then this fixed point is unstable. If all eigenvalues maintain the property, $|\lambda| < 1$, then the fixed point is stable. If $\lambda_\pm \in \mathbb{C}$ and $|\lambda_\pm| = 1$, then we say that the fixed point is neutrally stable and we see oscillations in solutions in its neighborhood. Lastly, if $\lambda_+ = \lambda_- = 1$, then the fixed point is unstable unless the initial conditions are rigged so that the second "guessed" solution vanishes. It turns out that these results are enough to derive explicit formulae for the problems our gambler faced earlier.

[11] Recall that the analogous statement from ordinary differential equations would be $y(t) = c_1 e^{\lambda_+ t} + c_2 e^{\lambda_- t}$ and growth and decay is determined by the **sign** of the corresponding eigenvalues with general asymptotic stability if both eigenvalues are negative. In the discrete case, it is the **size** of the eigenvalues that determines stability.

Problems:

1. ✎ Suppose a gambler bets \$1 and has a probability of p to win \$1. Let P_n be the probability that the gambler goes home a winner at a step with the current bankroll of \$$n$. Use the law of total probability to define a linear recurrence relation for this gambler's time series. Hint: At this step, we have $P_n = \alpha P_{n+1} + \beta P_{n-1}$, what should the values α and β be?

2. ✎ Assume that $P_n = \lambda^n$ and show that the recurrence relation defines the characteristic polynomial $p\lambda^2 - \lambda + (1-p) = 0$ and that this gives the roots $\lambda = 1$ and $\lambda = (1-p)/p$.

3. ✎ If $p = 1/2$, then we have repeated roots and $P_n = c_1 + c_2 n$. Apply the boundary conditions $P_0 = 0$ and $P_w = 1$ to arrive at $P_n = n/w$.

4. ✎ If $p \neq 1/2$, then we have distinct roots and

$$P_n = c_1 \left(\frac{1-p}{p}\right)^n + c_2 1^n. \tag{36}$$

Apply the boundary conditions to show that

$$c_1 = \frac{1}{\left(\frac{1-p}{p}\right)^w - 1}, \tag{37}$$

$$c_2 = \frac{-1}{\left(\frac{1-p}{p}\right)^w - 1}, \tag{38}$$

which reduces the solution to the recurrence relation to

$$P_n = \frac{\left(\frac{1-p}{p}\right)^n - 1}{\left(\frac{1-p}{p}\right)^w - 1}. \tag{39}$$

What is the meaning of this result?

5. ✎ Assume that $n \gg 1$ and justify that we have the following approximation,

$$P_n = \left(\frac{1-p}{p}\right)^{-m}. \tag{40}$$

Evaluate this approximation when \$$m$ = \$100 and explain what the result is trying to say about the gambler's situation.

6. ✎ How are the results different if the game is fair?

7. ✎ If H_n is the expected number of bets before ending the game by being broke or a winner while holding \$$n$, then $H_0 = 0$ and $H_w = 0$. Explain the meaning of the recurrence relation $H_n = 1 + pH_{n+1} + (1-p)H_{n-1}$.

8. ✍ The homogeneous recurrence relation is the same as before, and so we need only find the particular solution. Since the inhomogeneous term is constant, we guess $H_n = A$. Show that this doesn't work. We must then guess $H_n = nA$. Show that this assumption leads to $A = 1/(1-2p)$.

9. ✍ With our homogeneous and particular solutions in hand, we form the general solution to the recurrence relation,

$$H_n = c_1 \left(\frac{1-p}{p}\right)^n + c_2 + \frac{n}{1-2p}. \tag{41}$$

Apply the boundary conditions to arrive at,

$$c_1 = \frac{-\left(\frac{w}{1-2p}\right)}{\left(\frac{1-p}{p}\right)^w - 1}, \tag{42}$$

$$c_2 = \frac{\left(\frac{w}{1-2p}\right)}{\left(\frac{1-p}{p}\right)^w - 1}. \tag{43}$$

10. ✍ Using the previous results, simplify our expectation to

$$H_n = \frac{n}{1-2p} - \frac{w}{1-2p} P_n. \tag{44}$$

11. ✍ Show that if the house has any advantage, then $P_n \ll 1$ for $m \gg 1$ and from this arrive at Eq. (17).

2.2 *Lotka–Volterra Recurrence Relations*

> If you ask any mathematician if in his mind he makes a distinction between the theories of elasticity and those of electrodynamics, he will tell you that he does not, because the types of differential equations he encounters, and the methods which he must employ to solve the problems which arise, are all the same in the two cases.
>
> Vito Volterra (1860-1940)

OFTEN WE THINK ABOUT OUR RECURRENCE RELATION AS DISCRETE-TIME ORDINARY DIFFERENTIAL EQUATIONS. When we do, it is typical to call them difference equations. The Lotka-Volterra difference equations are given by,

$$x_{n+1} = f(x_n, y_n) = r x_n - \alpha x_n y_n, \tag{45}$$

$$y_{n+1} = g(x_n, y_n) = s y_n + \beta x_n y_n, \tag{46}$$

where $n \in \mathbb{N}_0 = \{0, 1, 2, \dots\}$ and $r, s, \in \mathbb{R}$ define the growth rates for the two species while, $\alpha, \beta \in \mathbb{R}$ characterize the nature of their interactions. The continuous version was first proposed by Lotka in 1920 to model the dynamics of an herbivorous animal species and the plant species it consumed. In 1926, Volterra proposed a similar set of equations to study fish catches in the Adriatic sea. Just like its continuous counterpart, the discrete version can manifest periodic solutions and qualifies as a type of nonlinear oscillator.[12] Thus, our intuition from mechanical analogs tells us that coupled predator-prey systems should exhibit complicated behaviors ranging from emergent properties, i.e., large-scale patterns not necessarily predicted by individual dynamics, to chaos, i.e., long-term dynamics which are highly sensitive to state conditions. Here we consider a first-pass of analysis informed by our work with ordinary differential equations and multivariate calculus.

[12] Important here is that the parametric curves in the $(x, y)-$plane are not parameterized by the typical trigonometric functions. Not that, at this point, we would have expected such a kindness.

Problems:

1. Find the fixed points of the map.
2. Find the Jacobian matrix, associated with the map, centered about each fixed point.
3. Calculate the eigenvalues of these Jacobian matrices.
4. Determine values for r and s such that the trivial fixed point is stable.
5. Show that if both species grow when not interacting with each other, then the nontrivial fixed point is unstable.
6. Show that if in the absence of interactions, one species decays while the other grows, then the populations oscillate about the nontrivial fixed point.
7. Using the code `LotkaVolterra.R`, simulate these oscillatory dynamics with $x_0 = y_0 = 1.25$, $r = 1.0015$, $s = 0.9994$, $\alpha = 0.0006$, and $\beta = 0.00025$, for 7100 time-steps.

2.3 *On the logistics equation*

> Not only in research, but also in the everyday world of politics and economics, we would all be better off if more people realised that simple nonlinear systems do not necessarily possess simple dynamical properties.
>
> Robert May (1936-2020) in *Simple mathematical models with very complicated dynamics* (1976), Nature 261: 459–467

Nonlinear maps, even at first-order, can generate complicated and chaotic outcomes.[13] In fact, even the simplest nonlinear upgrade to an autonomous first-order linear map, i.e., the

[13] This is in stark contrast to its continuous-time counterpart, which required three degrees of freedom.

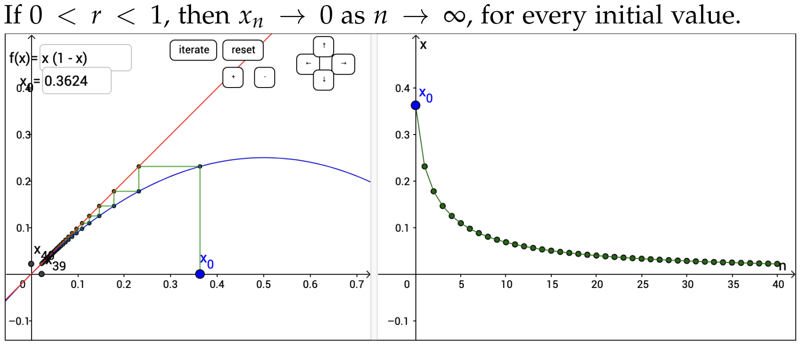
discrete logistics equation, makes for a rich test bed. As one should expect, a direct symbolic assault will not yield the results necessary for a decent understanding of the problem, and we will also need to leverage qualitative and numerical analyses. First, we consider the geometric analysis given to us by a cobweb plot. In a cobweb plot, we depict the dynamics as points in the (x_n, x_{n+1})-plane and pay special attention to fixed points, which may be stable or unstable, and correspond to intersections between the map and the $x_n = x_{n+1}$ line. At the same time, we must also be aware that autonomous first-order maps can generate sign oscillations where their continuous-time counterpart cannot. More astounding is that these oscillations can be pushed to a point where cycles, i.e., fixed points in the composition of maps, manifest. Here, there isn't much control we can place on how deep these limit cycles can appear in the repeated composition of maps, generating an attracting set on which trajectories accumulate. In fact, it is known that for a specific parameter regime, period-doubling bifurcations occur and that these period-doubling bifurcations increase so rapidly as the parameter is changed that chaos can result.

The logistics map is given by,

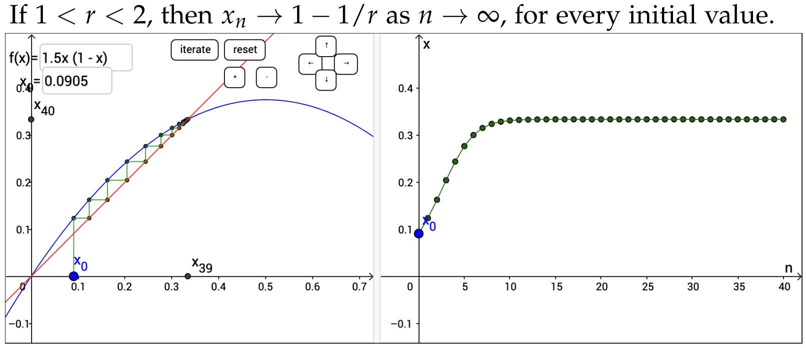
$$x_{n+1} = rx_n(1 - x_n), n = 1, 2, 3, \ldots . \tag{47}$$

If we define r_n to be the parameter value where a stable 2^n−cycle occurs, then we have the following progression:

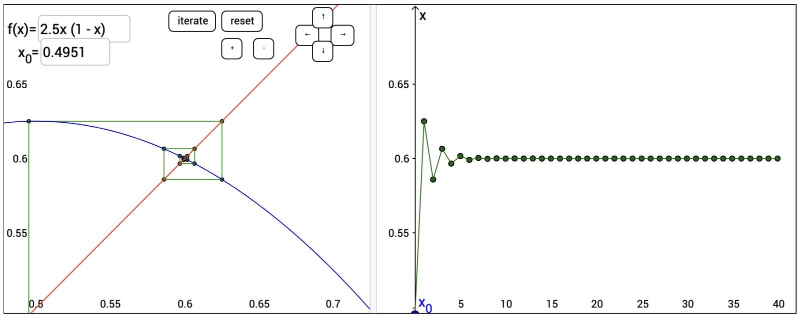
1. If $0 < r < 1$, then $x_n \to 0$ as $n \to \infty$, for every initial value.



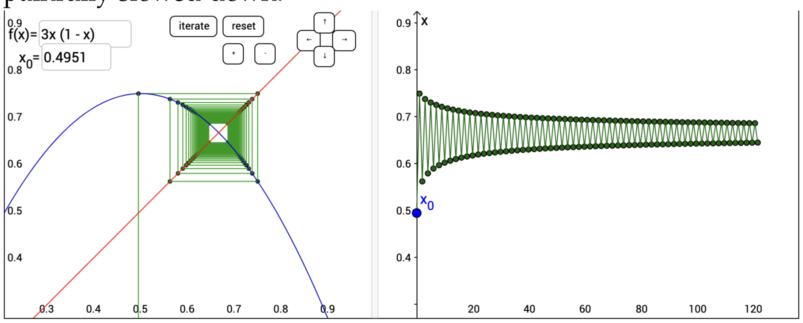
2. If $1 < r < 2$, then $x_n \to 1 - 1/r$ as $n \to \infty$, for every initial value.



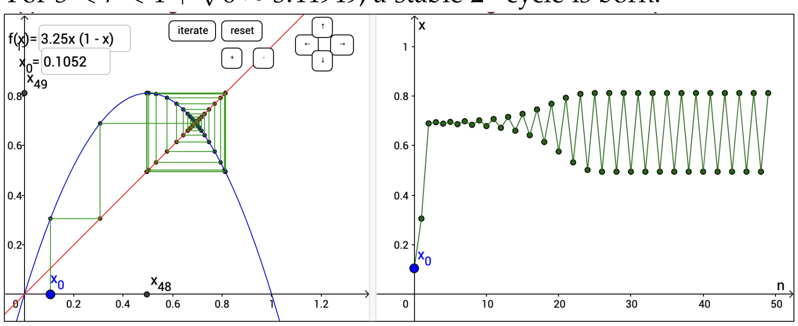
3. If $2 < r < 3$, then the same is true, but the system can oscillate as it does so.



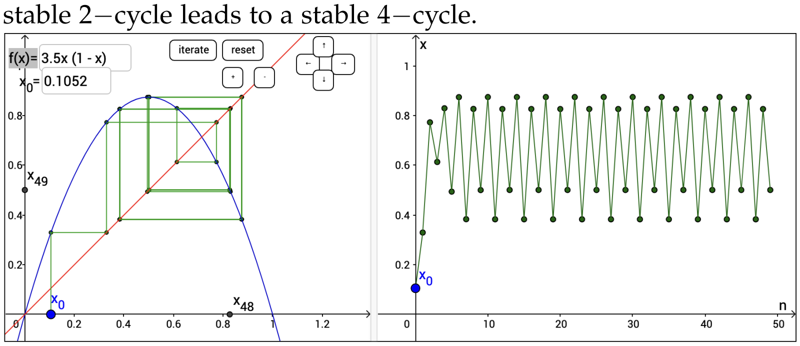
4. At $r = 3$, we still see convergence to the fixed point, except it is painfully slowed down.



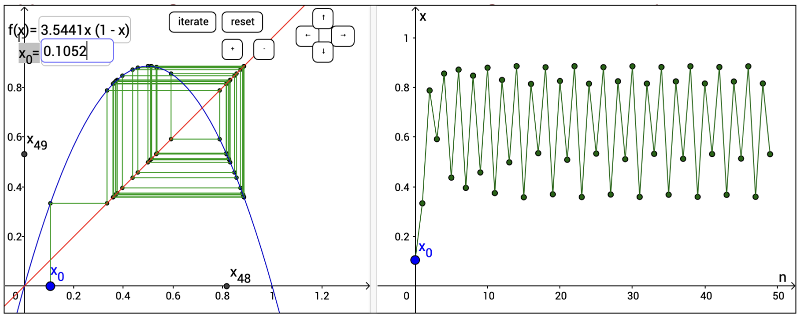
5. For $3 < r < 1 + \sqrt{6} \approx 3.44949$, a stable $2-$cycle is born.



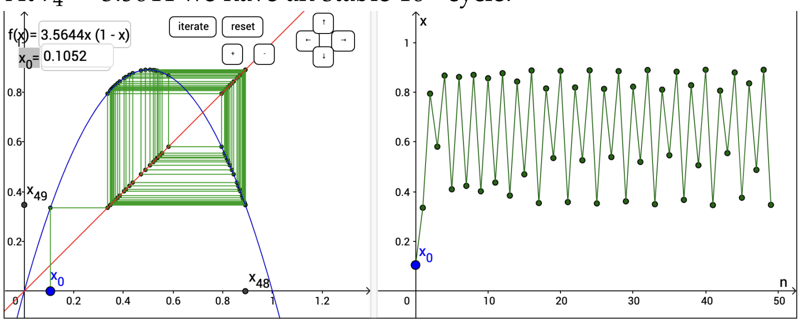
6. For $3.44949 < r < 3.54409$, a period-doubling bifurcation of the stable $2-$cycle leads to a stable $4-$cycle.



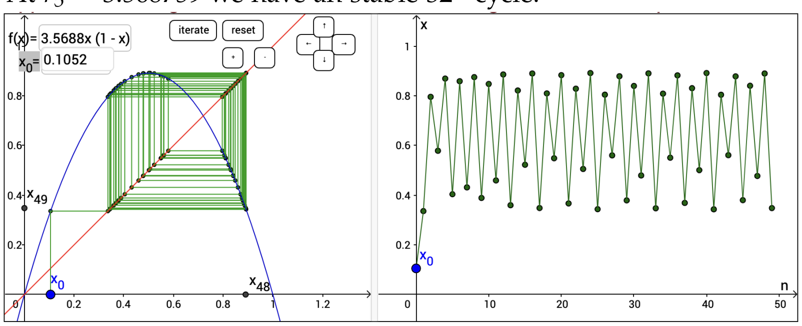
7. At $r_3 = 3.54409$ we have an stable $8-$cycle.



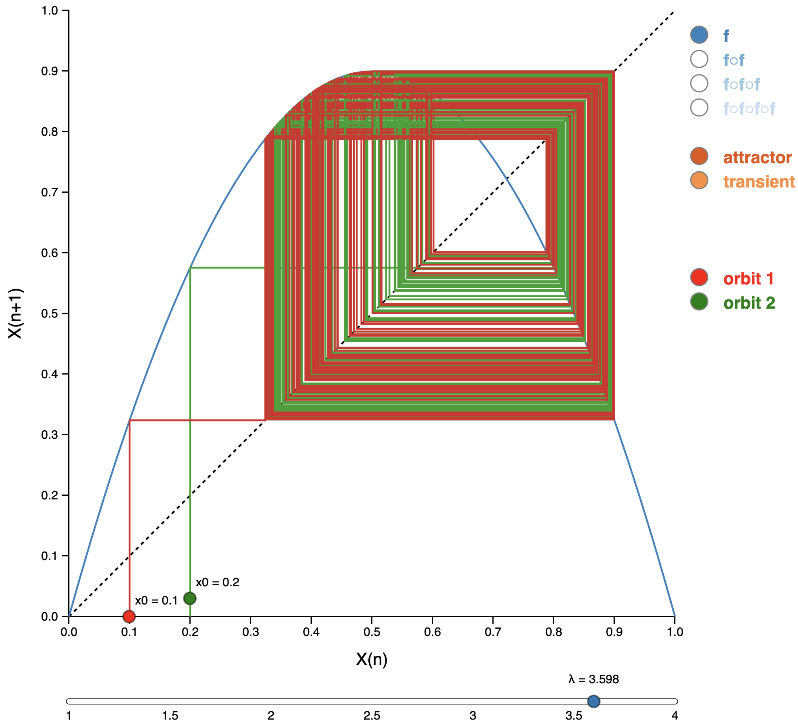
8. At $r_4 = 3.5644$ we have an stable $16-$cycle.



9. At $r_5 = 3.568759$ we have an stable $32-$cycle.



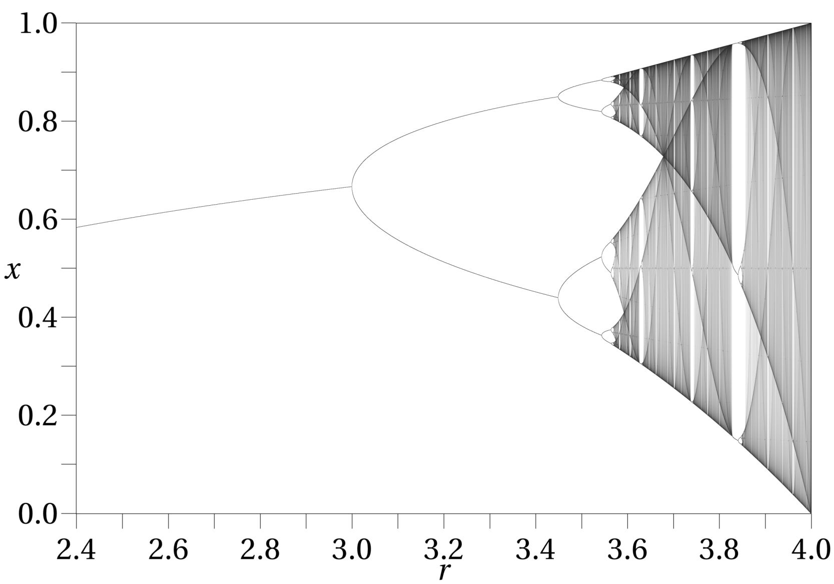
10. At this point, the period-doubling bifurcations occur with only minor changes to the parameter value, i.e., period-doubling cascade. The associated diagrams don't really hint at much of a change, but when two trajectories with similar initial values are plotted with each other, we can see very different behavior, which is a hallmark of dynamical chaos.



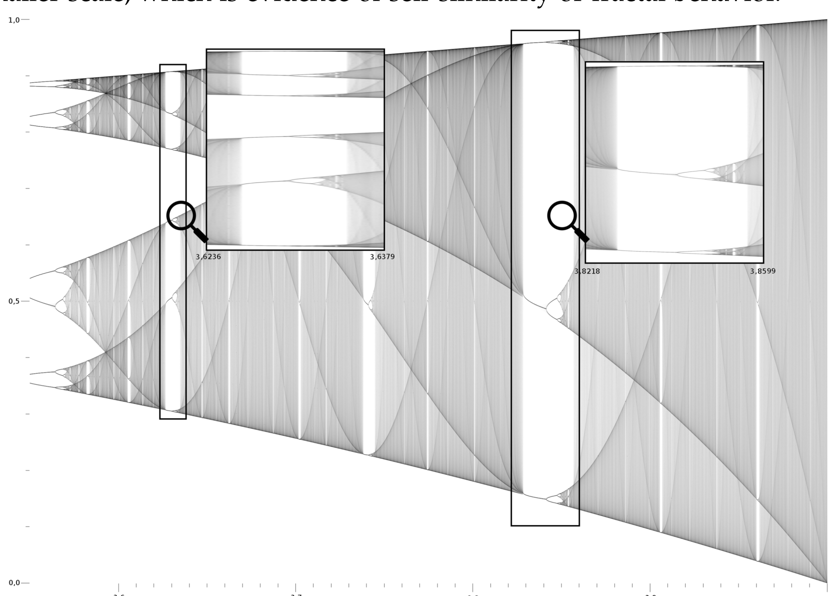
These higher-order cycles are produced as local minima/maxima of the composition of maps generate fixed points. It turns out that for chaos derived from a locally quadratic map, the ratio of successive r_n is a fixed number known as Feigenbaum's constant, $\delta = 4.699201609....$ That is, every chaotic system that corresponds to this description will bifurcate at the same rate![14]

While not technically the same as our continuous-time counterparts, we can plot the cycle points as the parameter is varied to generate a type of bifurcation diagram for the logistics map.

[14] Often we say that this constant is *universal* for such maps. It is suspected but not proven that Feigenbaum's constant is transcendental.



Interestingly, there exist intermittent stable regions of lower-order cycles. These regions of stability manifest as local extremes of higher-order compositions get close to the input-output line, and trajectories are forced to spend lots of time ricocheting around trying to pass through them. Interestingly, after these intermittent regions of low-order stable cycling, the whole process starts again, and if we zoom into these regions, we see a period-doubling cascade happening at a smaller scale, which is evidence of self-similarity or fractal behavior.



Problems:

1. Consider the applet at http://rocs.hu-berlin.de/D3/logistic/ and go through each of the regimes provided above. Write a one-paragraph explanation of what you are seeing to yourself at the start of the semester.

2. Watch 19.4 Dripping Faucet 8 14 2019, (`https://youtu.be/m_`

Z-SIxqYcI

3. 20.1 Chaos and Cardiac Arrhythmias https://youtu.be/Hvi-Hy0qLpg

4. This equation will change how you see the world (the logistic map) https://youtu.be/ovJcsL7vyrk

3 *Stochastic dynamics*

> Everything we care about lies somewhere in the middle, where pattern and randomness interlace.
>
> James Gleick (1954-present) in *The Information: A History, a Theory, a Flood*

3.1 *Logistics map with noise*

> Estimating is what you do when you don't know.
>
> Sherman Kent (1903-1986)

To be updated once we've had lecture topics.

3.2 *Stochastic discrete time Lotka-Volterra*

> Plans to protect air and water, wilderness and wildlife are in fact plans to protect man.
>
> Stewart Udall (1920-2010)

To be updated once we've had lecture topics.

3.3 *Introductory calculations for stochastic ordinary differential equations*

> We will always have STEM with us. Some things will drop out of the public eye and will go away, but there will always be science, engineering, and technology. And there will always, always be mathematics.
>
> Katherine Johnson (1918-2020)

To be updated once we've had lecture topics.