last revised: 30 October 2009

14 The Dynamics of Balance Theory

In some types of social networks, the links between actors can be either positive or negative. Depending on the context, positive links might reflect friendship or alliance; negative links might reflect enmity or hostility. *Balance theory* attempts to explain the structure of these networks. After briefly reviewing some core ideas of balance theory, we develop a simple model of the adjustment process by which links might change sign (from positive to negative or vice versa). The model is technically an absorbing Markov chain with 2^N states (where N is the number of links). However, as we verify through simulation analysis, the behavior of this Markov chain is well approximated by a simple one-dimensional nonlinear model. Interestingly, given the myopic adjustment process we consider, the equilibrium of the nonlinear model may differ from the long-run outcome ("balance") presumed by the theory.

14.1 Balance theory

In his seminal contribution to balance theory, Heider (1946) focused on a simple social network with three actors (i.e., a *triad*) in which each pair of actors is connected by either a positive or negative link. Graphically, we may use solid lines to denote positive links, and dotted lines to denote negative links. Ignoring the identities of the actors, there are four possible types of triads, shown below.

0 **1** **2** **3**

Note that the index for each triad simply indicates the number of negative links. Heider argued that, from a social-psychological perspective, triads of type 0 or 2 are stable ("balanced") while triads of type 1 or 3 are not. Essentially, type-1 triads are imbalanced because they violate the principle that "a friend of my friend is my friend." If the network started as a type-1 triad, balance might be achieved either by making the negative link positive (so that the network becomes a type-0 triad) or by making one of the positive links negative (so that the network becomes a type-2 triad). Essentially, type-3 triads are imbalanced because they violate the principle that "an enemy of my enemy is my friend." If the network started as a type-3 triad, we might thus anticipate that one of the negative links will become positive, so that the network is transformed into a type-2 triad. In contrast, because triads of type 0 and 2 satisfy both of these principles, we have no reason to expect any change in the sign of the links. Given enough time for social-psychological forces to "play out," we might thus expect to observe only balanced triads (type 0 or 2) in the long run.[1]

In another seminal paper, Cartwright and Harary (1956) formalized Heider's intuition, and extended the concept of balance to larger networks. Interestingly, their concept of balance can be stated in two equivalent ways. The first requires some terminology from graph theory. A *cycle* is a sequence of links that begins and ends at the same node. For a signed graph (i.e., a network in which links are either positive or negative), the *sign of a cycle* is given by the product of the signs of the links that comprise the cycle. To illustrate, consider again the 4 types of triads. From the diagram above, it is apparent that type-0 triads contain a positive cycle (schematically, $+ \cdot + \cdot + = +$), type-1 triads contain a negative cycle ($+ \cdot + \cdot - = -$), type-2 triads contain a positive cycle ($+ \cdot - \cdot - = +$), and type-3 triads contain a negative cycle ($- \cdot - \cdot - = -$). This suggests that the balance of a signed graph depends on the signs of its cycles. More precisely, following the definition developed by Cartwright and Harary, a signed graph is balanced if and only if every cycle is positive. Importantly, this definition allows us to move beyond triads, assessing balance for social networks with any number of actors.[2]

Balance can also be assessed in a second way. Again following Cartwright and Harary (1956), a signed graph is balanced if and only if the nodes can be partitioned into two subsets (one possibly empty), with any positive links connecting nodes within subsets, and any negative links connecting nodes between subsets. Intuitively, the network is balanced when the actors can be separated into (no more than) two "camps" in such a way that any friends belong to the same camp and any enemies belong to different camps. Returning to our 4 types of triads, note that this partition is possible for triads of type 0 (where all actors belong to the same camp) and type 2 (where 2 actors belong to one camp and the remaining actor belongs to the other camp). However, this partition is not possible for triads of type 1 or 3.[3] Like the "cycle" criterion discussed above, the "partition" criterion can also be used to assess balance in networks with any number of actors.[4]

[1]Against Heider (1946), some researchers (notably Davis 1967) have argued that type-3 triads are not inherently unstable. This leads to a variation on balance called "clusterability." For the present chapter, we focus on Heider's original conception of balance.

[2]For a signed graph with many nodes, some cycles may be quite long. This makes it difficult to assess balance merely by inspecting the graph. However, researchers have developed algebraic procedures for assessing balance (see Harary et al 1965, Batagelj 1994, Montgomery 2009).

[3]For the type-1 triad, suppose we label the actors $\{A, B, C\}$ so that there is a positive link from A to B, a positive link from B to C, and a negative link from A to C. By virtue of the positive links, both A and C should be placed in B's camp. But the negative link from A to C makes this impossible. For the type-3 triad, we can eliminate negative links within camps only if there are 3 different camps. But balance permits at most 2 camps.

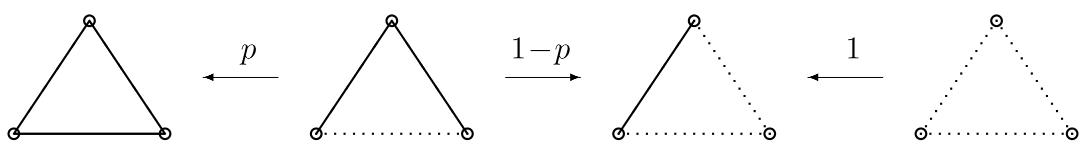
[4]While the equivalence of the cycle and partition criteria may not be immediately obvious, this is

Cartwright and Harary (1956) did not presume that the network is complete (i.e., that every pair of actors is connected by either a positive or negative link). However, if we restrict attention to complete networks, the balance criterion may be stated in yet another way. A complete network with n actors has $n(n-1)/2$ links and $n(n-1)(n-2)/6$ triads. As first noted by Flament (1963), a complete signed graph is balanced if and only if every triad is balanced. To assess balance using this "triad" criterion, we can conduct a *triad census*, counting the number of triads of each of the 4 possible types. The network is balanced if and only if every triad is of type 0 or 2 (and hence no triads are of type 1 or 3). Moreover, a simple measure of the "degree" of balance is given by the proportion of triads that are balanced.

14.2 A process model

Given that background, we might view "balance" as a conjecture about the structure of the network in the long run. However, without a specification of the process by which links are altered, it is unclear whether this outcome will actually be reached. While researchers have proposed a variety of adjustment mechanisms, this section considers a very simple "myopic" specification from Antal et al (2006).

For simplicity, we assume a complete network, so that each pair of actors is connected by either a positive or negative link. Each period, we randomly select one triad. If this triad is balanced (type 0 or 2), no links are altered. If the triad is type 3, then one of the negative links (chosen at random) becomes positive. If the triad is type 1, then the negative link becomes positive with probability p. Otherwise (with probability $1-p$), one of the positive links (chosen at random) becomes negative. Our adjustment process is thus summarized by the following diagram.



0 **1** **2** **3**

Note that, in the special case where $p = 1/3$, we simply "flip" the sign of a random edge whenever an unbalanced triad is selected. While this adjustment process ensures balance of the focal triad, the process is "myopic" because any sign change could induce imbalance in other triads. After modifying the focal triad (if it was imbalanced), we begin the next period by randomly selecting a new focal triad. This

the content of the *Structure Theorem* proven by Cartwright and Harary (1956). Interested readers might see the original paper or Chartrand (1977) for a proof.

adjustment process is repeated indefinitely (unless balance is achieved, in which case no further adjustments occur).

Formally, this process constitutes a Markov chain. The states of the chain are the possible configurations of the network. Thus, given n actors, there are $2^{n(n-1)/2}$ possible states of the chain. If the number of actors is very small (say $n = 3$ or 4), we could analyze this model using the (transition matrix) method developed in Chapter 4. Given larger n, this approach is clearly impractical. However, based on our knowledge of Markov chains, we can still say something qualitative about the long-run behavior of this chain. From the description of the adjustment process, it is apparent that balanced states (in which all triads are positive) are absorbing, that imbalanced states (in which some triads are negative) are transitory, and that every imbalanced state can reach (directly or indirectly) a balanced state. Thus, we know that the chain will eventually be absorbed if the process continues indefinitely.

While that conclusion is correct, it is also potentially misleading. Simulations reveal that, when n is large and $p < 1/2$, balance is unlikely to occur within any "empirically relevant" time frame.[5] Rather, over the course of a (very long) simulation run, the system appears to converge to a (stochastic) steady state with a constant proportion of positive links, and a degree of balance below 1. To illustrate, we use the m-file `balancedynamics` placed in Appendix xx.

```
>> help balancedynamics

  [d, y] = balancedynamics(p, x0, n, T)
  balance theory simulation model based on Antal et al (2006)
  input p = probability that type-1 triad becomes type-0 triad
  input x0 = probability that each link is initially positive
  input n = number of actors
  input T = number of iterations
  output d is a T x 4 matrix giving triad census [d0 d1 d2 d3] for each period
  output y is a vector of length T giving number of positive links for each period
```

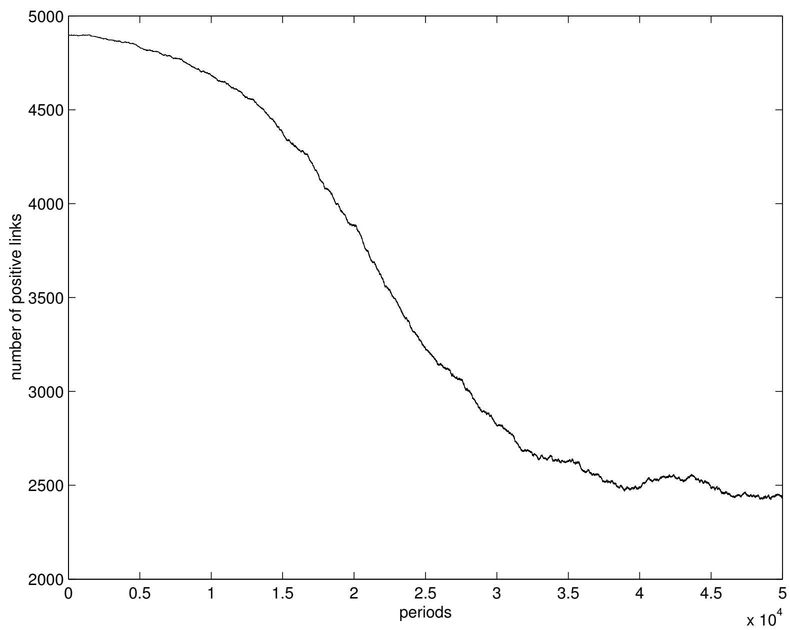
This program begins by constructing a random graph with n nodes (and hence $n(n-1)/2$ links) in which each link is positive with probability x_0 (and hence negative with probability x_0). It then implements (for T iterations) the adjustment process described above so that a type-1 triad becomes a type-0 triad with probability p (and hence becomes a type-2 triad with probability $1 - p$). The function outputs a $(T \times 4)$ matrix d giving the triad census for each period, and a (length T) vector y giving the number of positive links in each period.

Setting the parameters $p = 1/3$, $x_0 = .99$, $n = 100$, and $T = 50000$, the results of one trial are shown on the next page.[6] For these parameters, the total number of links

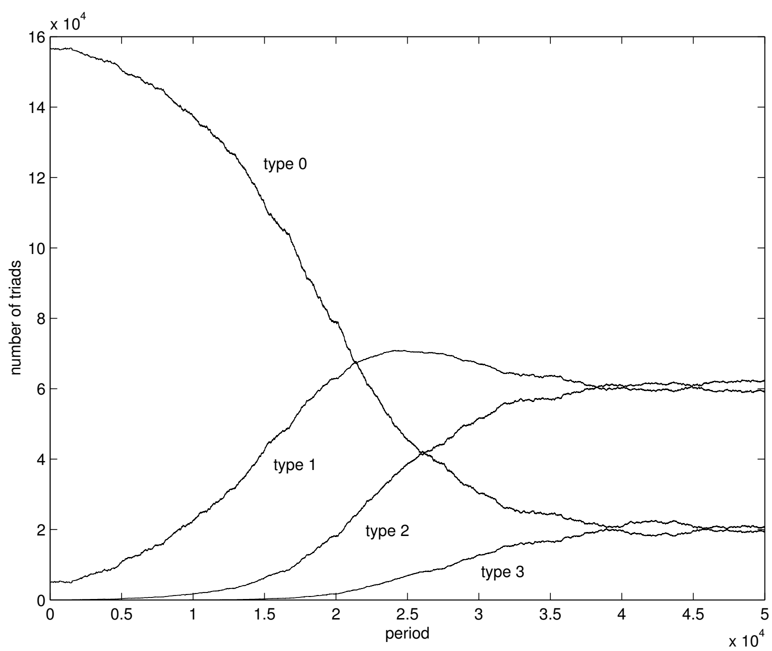
[5] See Antal et al (2006) for some formal analysis of the relationship between network size and expected time to absorption.

[6] On my desktop computer, this run took about 10 minutes. The long computation time is due

```
>> [d, y] = balancedynamics(1/3, .99, 100, 50000);
>> plot(y)  % number of positive links over simulation run
```




```
>> plot(d)   % triad census over simulation run
```



is $n(n-1)/2 = 4950$, and the total number of triads is $n(n-1)(n-2)/6 = 161700$. Thus, from the top diagram, we see that the share of positive links falls from its initial value (approximately $x_0 = 0.95$) to a "steady state" share of approximately $x^* = 1/2$. From the bottom diagram, we see that the shares of type-1 and type-2 triads converge to approximately 3/8, while the shares of type-0 and type-3 triads converge to approximately 1/8.[7]

Further simulation runs produce quite similar results. (Readers are encouraged to use the m-file to verify this claim for themselves. In particular, you should vary x_0 to demonstrate that the same imbalanced equilibrium arises for any initial condition.) Thus, for our current parameter values ($p = 1/3$ and n large), simulation analysis suggests that the degree of balance (i.e., the sum of the shares of type-0 and type-2 triads) does not approach 1, but instead converges to 1/2. Of course, if we allowed this simulation to continue forever, we know that the chain will eventually reach a balanced (absorbing) state. But based on even longer simulation runs (and analytical results from Antal et al 2006), it is apparent that the system will "hover around" an imbalanced equilibrium for any "empirically relevant" time frame.

14.3 A deterministic approximation

The smooth time paths shown in the preceding diagrams might suggest the possibility of approximating the (very high-dimensional) Markov chain process with a (hopefully low-dimensional) nonlinear system. In fact, Antal et al (2006) have shown that the behavior of the process is well approximated by a one-dimensional nonlinear system. To begin developing this model, let x_t denote the proportion of links that are positive in period t. Assuming that process begins with a random graph (so that the probability of a positive realization is independent across links), we can obtain a good approximation of the triad census by viewing each triad as a sample of 3 links. Following Antal et al (2006), we may say that the elements of the triad census are initially "uncorrelated." Interestingly, some formal analysis in Antal et al (2006) reveals that, if the elements of the triad census are initially uncorrelated, then they remain uncorrelated indefinitely. Consequently, for any period, the triad census is well approximated by the equations

$$\begin{aligned} d_t(0) &= {x_t}^3 \\ d_t(1) &= 3{x_t}^2(1 - x_t) \end{aligned}$$

primarily to recomputation of the triad census at every iteration. Computation time rises linearly in T and exponentially in n.

[7] The share of positive links (top diagram) can be derived from the triad census (bottom diagram) using the equation

$$x = \frac{3d(0) + 2d(1) + d(2)}{n - 2}$$

where $d(i)$ indicates the number of triads of type i. Intuitively, each type-i triad contains $3 - i$ positive links, and each link appears in $n - 2$ triads.

$$\begin{aligned} d_t(2) &= 3x_t(1-x_t)^2 \\ d_t(3) &= (1-x_t)^3 \end{aligned}$$

where $d_t(i)$ denotes the proportion of triads of type i in period t. Intuitively, viewing a triad as a sample of 3 links, we obtain a type-i triad when $3-i$ positive links and i negative links are drawn.

Because each element of the triad census depends solely on the proportion of positive links (x), we merely need to account for the dynamics of this one state variable. To develop an equation, consider one iteration of the process model. If a type-1 triad is selected (this occurs with probability $d(1)$) then the number of positive links rises by 1 with probability p and falls by 1 with probability $1-p$. If a type-3 triad is selected (this occurs with probability $d(3)$) then the number of positive links rises by one. Thus, the expected change in the number of positive links is given by

$$\begin{aligned} & d(1)[p-(1-p)] + d(3) \\ = \; & 3x^2(1-x)(2p-1) + (1-x)^3 \end{aligned}$$

Because x denotes the proportion (not number) of links that are positive, we need to divide this expression by the number of links in order to obtain the change in x. Doing so, we obtain our one-dimensional system

$$\Delta x \; = \; \frac{1}{n(n-1)/2} \; \left[3x^2(1-x)(2p-1) + (1-x)^3\right]$$

with parameters p and n.

As a first "test" of this approximation, we compute the predicted time paths for our preceding example, which are shown on the next page. Further examples confirm that there is a good match between simulation results and this one-dimensional model. (Readers are encouraged to compare simulation results to the deterministic predictions for different values of the parameter p and the initial condition x_0. You might also experiment with different values of n, though computation time increases exponentially in that parameter.)

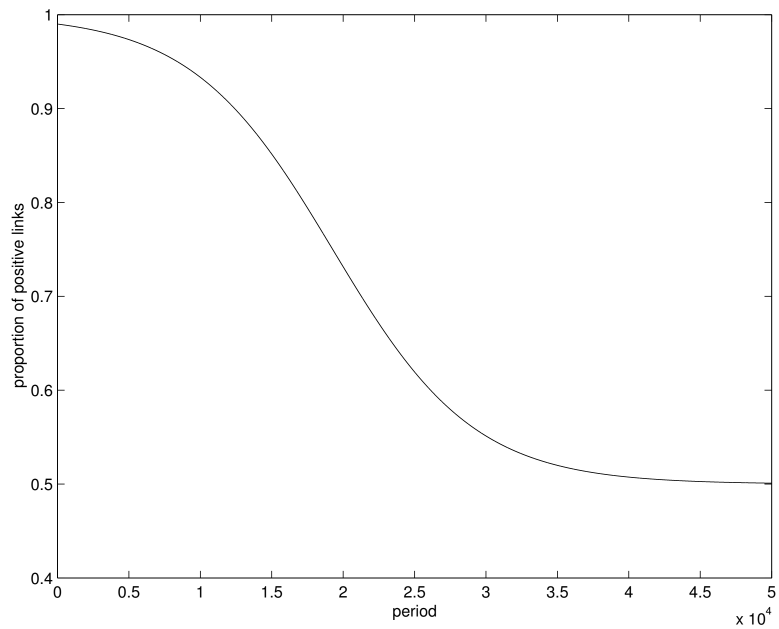
Returning to the equation above, it is apparent that Δx is always positive when $p > 1/2$ and $x \neq 1$. That is, when type-1 triads are more likely to be transformed into type-0 triads than type-2 triads, the proportion of positive links will continue to rise until all links are positive. Thus, given $p > 1/2$, we find that the network does become balanced (within an "empirically relevant" time frame). In contrast, for $p < 1/2$, we see that $\Delta x = 0$ implies $x = 1$ or

$$3x^2(2p-1) + (1-x)^2 \; = \; 0$$

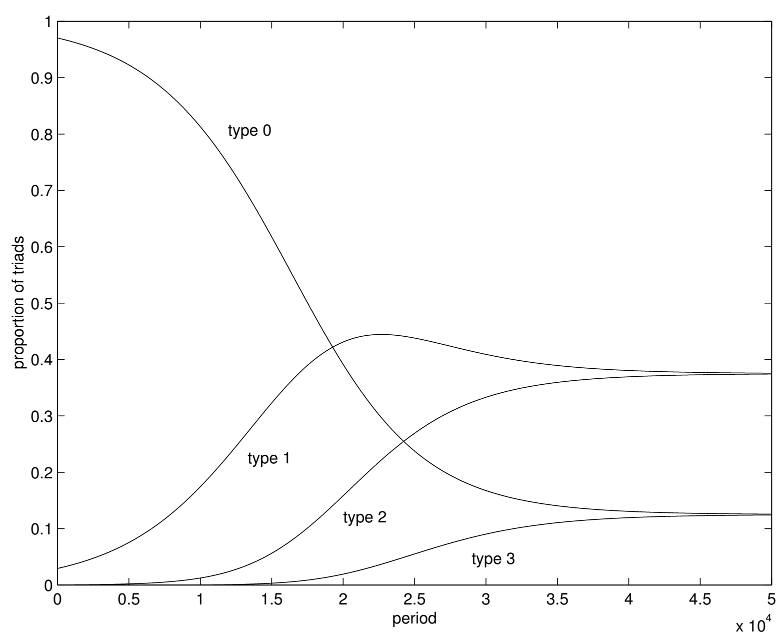
Solving this quadratic equation, we obtain

$$x \; = \; \frac{2 \pm \sqrt{4 - 4(6p-2)}}{2(6p-2)} \; = \; \frac{1 \pm \sqrt{\delta}}{1-\delta} \; = \; \frac{1}{1 \pm \sqrt{\delta}}$$

```
>> x = .99; p = 1/3; n = 100; L = n*(n-1)/2; y = x;
for t = 1:50000; dx = (3*x^2*(1-x)*(2*p-1) + (1-x)^3)/L; x = x+dx; y = [y; x]; end
>> plot(y)   % time path for proportion of positive links
```




```
>> plot([y.^3, 3 * y.^2 .* (1-y), 3 * y .* (1-y).^2, (1-y).^3])
>> % time paths of triad census
```



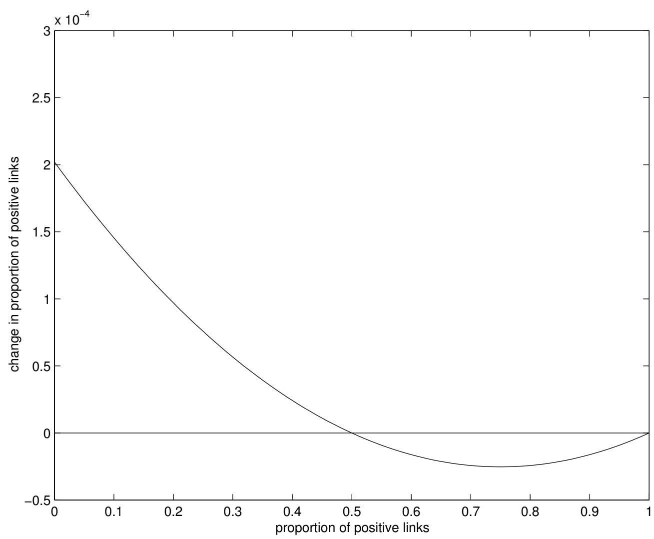
where $\delta = 3 - 6p$. Focusing on the root that lies between 0 and 1, we thus obtain the steady state

$$x^* = \frac{1}{1 + \sqrt{3 - 6p}}$$

Because Δx is positive for x between 0 and x^*, and is negative for x between x^* and 1, the steady state x^* is the unique stable equilibrium.

Fixing a particular value of p, we could have found this steady state graphically by plotting Δx as a function of x, with x^* determined by the intersection of Δx with the horizontal axis. To illustrate, we suppose that $p = 1/3$, and plot Δx below.

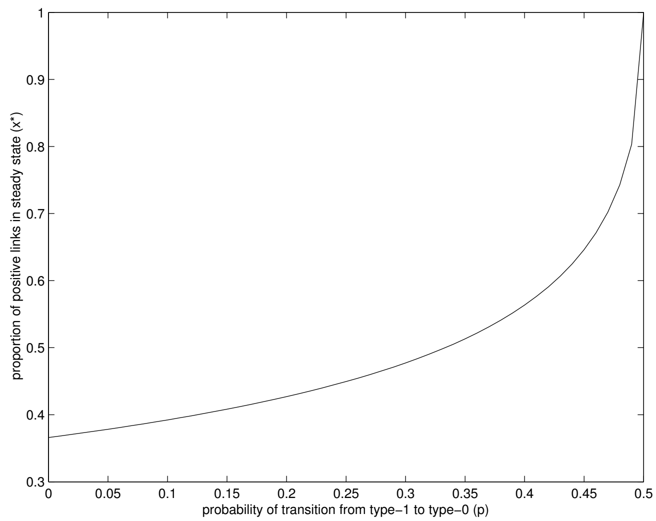
```
>> p = 1/3; n = 100; L = n*(n-1)/2; x = 0:.01:1;
y = (3 * x.^2 .* (1-x) * (2*p-1) + (1-x).^3)/L;
plot(x,y,[0 1],[0 0])   % plot delta x as a function of x
```



As we've already learned from the time path diagram, this diagram indicates that the system converges to $x^* = 1/2$ when $p = 1/3$. We also learn from this diagram (and could verify through simulation) that the process approaches the steady state more rapidly from below (when $x_0 < x^*$) than above (when $x_0 > x^*$).

Alternatively, we can plot x^* as a function of p to see how the steady state changes as we vary this adjustment parameter.

```
>> p = 0:.01:.5; x = 1./(1+sqrt(3-6*p))
>> plot(p,x) % steady state x as a function of transition probability p
```



From this diagram, we may once again confirm that $p = 1/3$ implies $x^* = 1/2$. But we also learn that p has a nonlinear effect on x^*. More precisely, an increase in p has a relatively small effect on x^* when p is close to 0. But when p is near 1/2, small increases in p have a large effect on the steady state x^*.

14.4 Further reading

We have already noted the important contributions made to balance theory by Heider (*J Psych* 1946) and Cartwright and Harary (*Psych Rev* 1956). The graph-theory textbook by Harary, Norman, and Cartwright (*Structural Models*, Wiley, 1965) offers a useful review. Chartrand (*Introductory Graph Theory*, Dover, 1977) provides an elementary introduction, and also discusses clusterability (introduced by Davis, *Human Relations* 1967). Early attempts to develop a process model were made by Flament (*Applications of Graph Theory to Group Structure*, Prentice-Hall, 1963) and Abell (*Sociology* 1968). Recent sociological contributions include Doreian and Krackhardt (*J Math Soc* 2001) and Hummon and Doreian (*J Math Soc* 2003). The adjustment process considered in this chapter, along with one-dimensional nonlinear approximation, are drawn from Antal et al (*Physica D* 2006). That paper also considers a second (less myopic) adjustment rule which causes networks to rapidly become balanced.

14.5 Appendix

14.5.1 Balancedynamics m-file

```
function [d, y] = balancedynamics(p, x0, n, T)
% [d, y] = balancedynamics(p, x0, n, T)
% balance theory simulation model based on Antal et al (2006)
% input p = probability that type-1 triad becomes type-0 triad
% input x0 = probability that each link is initially positive
% input n = number of actors
% input T = number of iterations
% output d is a T x 4 matrix giving triad census [d0 d1 d2 d3] for each period
% output y is a vector of length T giving number of positive links for each period

L = n*(n-1)/2;                      % number of links
A = rand(n) < x0; A = triu(A,1);    % initial random graph

y = []; d = [];
for t = 1:T

    y = [y; sum(sum(A))];  % count number of positive links

    P = A|A'; N = ~P - eye(n);
    d0 = trace(P*P*P)/6;  d1 = trace(P*P*N)/2; d2 = trace(P*N*N)/2; d3 = trace(N*N*N)/6;
    d = [d; d0 d1 d2 d3];   % triad census

    v = randperm(n); v = v(1:3); v = sort(v);
    s12 = A(v(1),v(2)); s23 = A(v(2),v(3)); s13 = A(v(1),v(3)); s = s12 + s23 + s13;

    if s == 1 | s == 3   % triad is balanced
        continue
    end

    r = rand;

    if s == 0           % all edges are negative
        if r < 1/3
            A(v(1),v(2)) = 1;
        elseif r < 2/3
            A(v(2),v(3)) = 1;
        else
            A(v(1),v(3)) = 1;
        end
    end

    if s == 2           % one edge is negative
        if s12 == 0
            if r < .5*(1-p)
                A(v(2),v(3)) = 0;
            elseif r < (1-p)
                A(v(1),v(3)) = 0;
```

```
            else
                A(v(1),v(2)) = 1;
            end
        elseif s23 == 0
            if r < .5*(1-p)
                A(v(1),v(2)) = 0;
            elseif r < (1-p)
                A(v(1),v(3)) = 0;
            else
                A(v(2),v(3)) = 1;
            end
        else
            if r < .5*(1-p)
                A(v(1),v(2)) = 0;
            elseif r < (1-p)
                A(v(2),v(3)) = 0;
            else
                A(v(1),v(3)) = 1;
            end
        end
    end
end
```