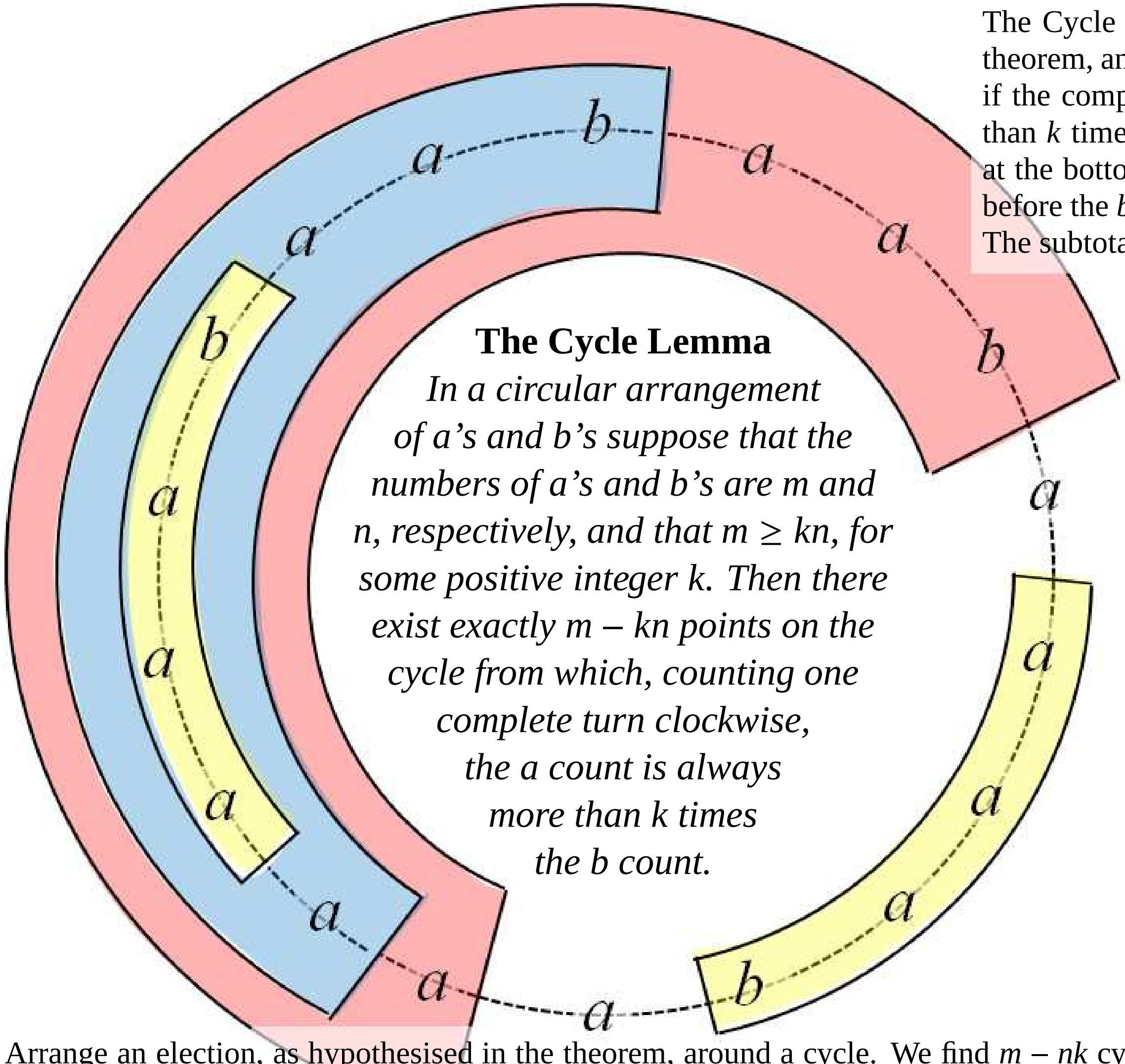
THEOREM OF THE DAY

Bertrand's Ballot Theorem *In an election where m people vote for candidate a and n for candidate b, suppose that $m > kn$, for some positive integer k. Then the probability that candidate a has always, from the first vote onwards, more than k times as many votes as candidate b is given by $(m - kn)/(m + n)$.*



The Cycle Lemma

In a circular arrangement of a's and b's suppose that the numbers of a's and b's are m and n, respectively, and that $m \geq kn$, for some positive integer k. Then there exist exactly $m - kn$ points on the cycle from which, counting one complete turn clockwise, the a count is always more than k times the b count.

The Cycle Lemma, below, is a clever way of explaining the numerator $m - kn$ in the theorem, and is useful in its own right. Suppose we say that a point on the cycle is 'good' if the complete clockwise turn of the cycle, starting at that point, always counts more than k times as many a's as b's. In the illustration on the left, take $k = 3$. Then the a at the bottom of the cycle is good, since we count 6 a's before the first b; then 2 more before the b at the top; another 2 before the penultimate b; and 4 more before the final b. The subtotals, 6, 8, 10, 14, are always more than 3 times the number of b's encountered.

The lemma says there are $m - kn$ good points. To see why it is true, write (for arbitrary m and n) $m = n(k - 1) + S$, for some positive integer S. We think of S as a 'surplus'. If $S = 0$ then the cycle may consist of n b's, each followed by $k-1$ a's. However, if $S \geq 1$ there must be at least one sequence on the cycle consisting of k a's followed by one b. The crucial observation is: no point in this sequence is good, whereas any point elsewhere is good if and only if it is good after removing the whole sequence (the b and k a's cancel each other out in our clockwise count). We can calculate that removing a sequence of k a's and one b gives a new cycle in which the surplus S is reduced by exactly 1. So provided $S \geq n$ to start with, we can repeatedly remove sequences of k a's followed by one b, until all n b's have been removed. The number of a's remaining will be $m - kn$: all of these a's are trivially good; they are therefore good in the original cycle.

Our illustration shows two 'inner' yellow sequences being removed (boxed), followed by a 'middle' blue sequence, and finally an 'outer' red sequence. The initial surplus is $S = 6$; after the yellow sequences are removed there remain 2 b's, 8 a's, and the surplus S' satisfies $8 = 2 \times 2 + S'$, so the surplus has reduced by 2, as expected. The $m - kn = 14 - 3 \times 4 = 2$ a's which finally remain are good points. And these correspond to cyclic permutations of a given sequence of votes for a and b which satisfy the conclusion of the Ballot Theorem.

Arrange an election, as hypothesised in the theorem, around a cycle. We find $m - nk$ cyclic rearrangements of the election satisfying its conclusion. This will overcount if a repeating pattern round the cycle produces identical rearrangements. For example: *aabbaaabaaaaabbaaabaaa*, with $k = 2$, yields four 'good' elections but only two distinct ones. However, the same periodicity means that 11 of all 22 cyclic rearrangements are distinct. Indeed, the fraction, say, α, of good elections which are distinct must always be the same as the fraction of all $m + n$ cyclic rearrangements which are distinct. So the probability that an election will be good is $\alpha(m - nk)/\alpha(m + n)$.

This theorem, in the case $k = 1$, is named for Joseph Bertrand, 1887; the Cycle Lemma is due to Aryeh Dvoretzky and Theodore Motzkin, 1947.

Web link: webspace.ship.edu/msrenault/ballotproblem/

Further reading: *Classic Problems of Probability* by Prakash Gorroochurn, Wiley-Blackwell, 2012.