Brownian motion within species and in phylogenies

Lecture 2-3, Tuesday, June 10

Joe Felsenstein

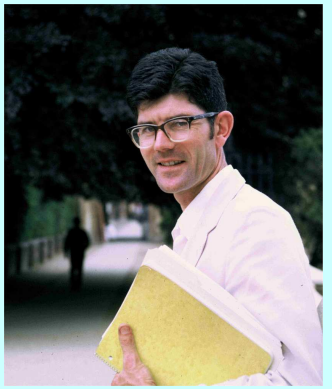
Evolutionary Quantitative Genetics Workshop 2025

What will approximate change of quantitative characters?

- ... when it occurs by genetic drift of pre-existing alleles?
- ... when it also occurs by mutation to new alleles?
- ... when variable selection (on othr characters) affects the alleles at each locus?
- ... when selection is on the fitness based on the whole phenotype?

Edwards and Cavalli-Sforza's approximation

Luca Cavalli-Sforza (and Edwards), 1963



Anthony Edwards, 1970

The expectation of gene frequency change in one generation (under pure genetic drift without mutation) is zero. The variance is the binomial variance

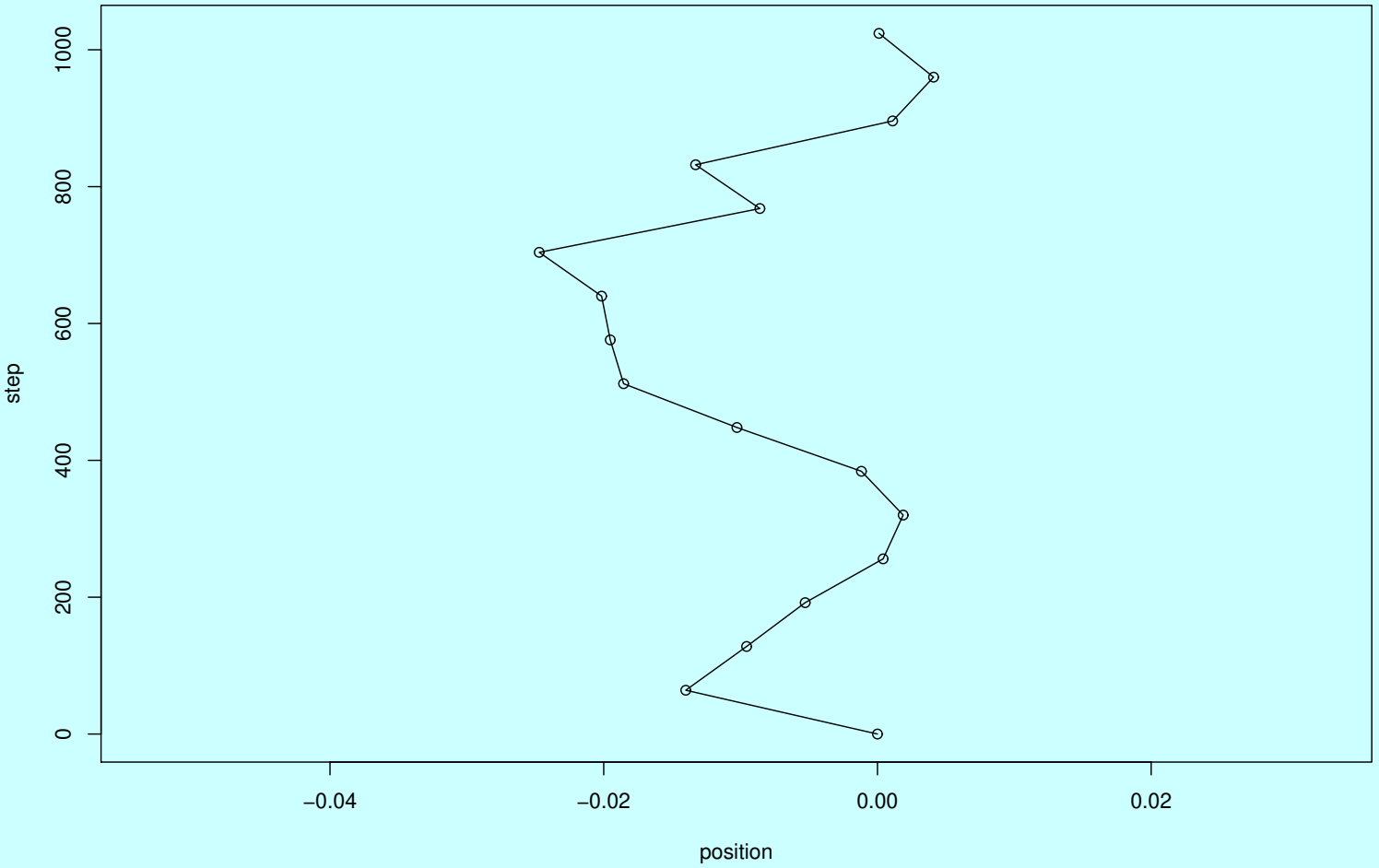
$$E\left[(\Delta p)^2\right] = \frac{p(1-p)}{2N_e}$$

That variance is not constant: it varies with p (in a parabola), but maybe we can roughly approximate it by dealing with the case where all populations have roughly similar gene frequencies, so the variances are nearly the same. Maybe. Roughly.

Brownian motion is mathematically tractable

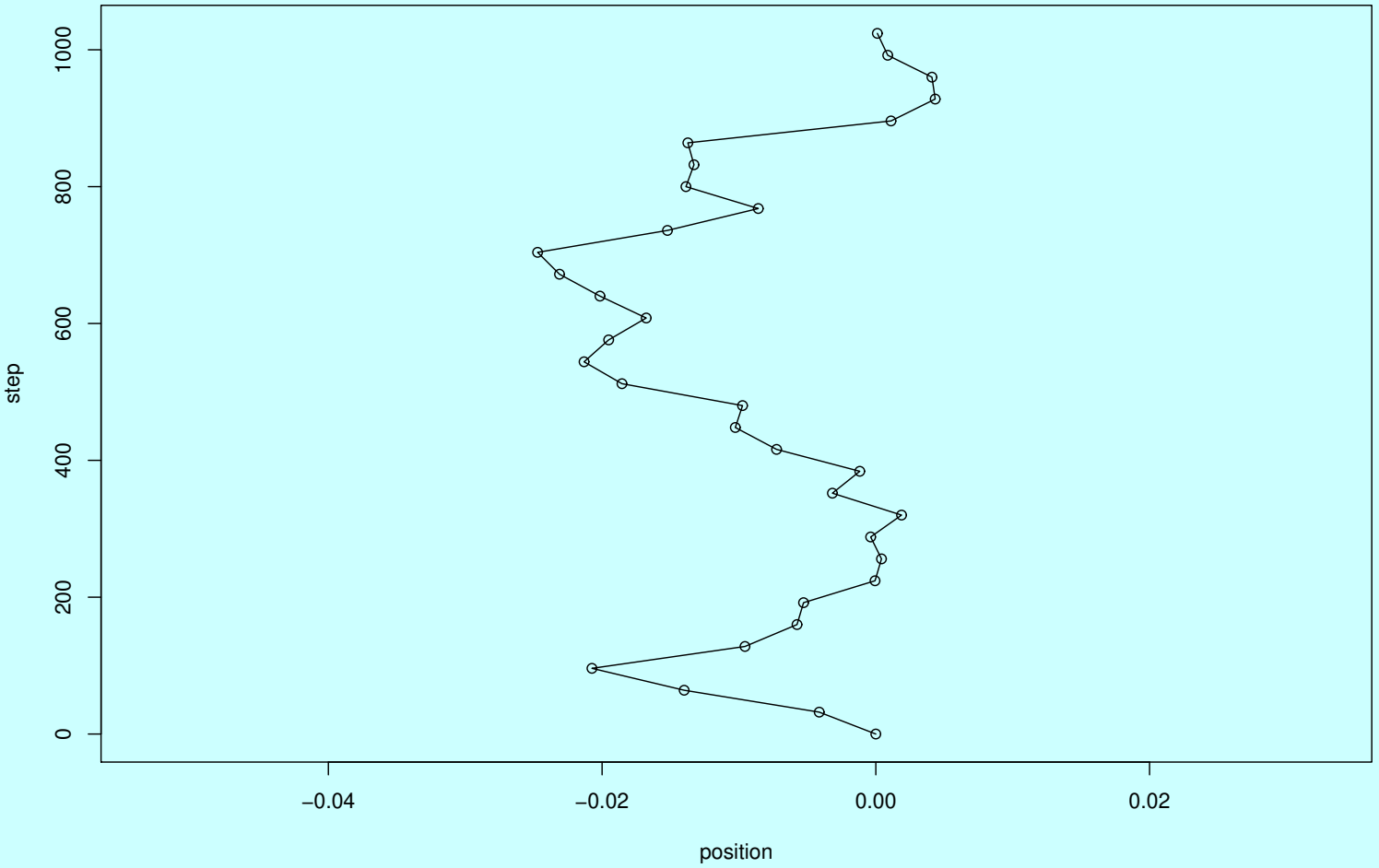
You can easily compute transition probabilities from one value to another, since the net change after "time" t is normal. with mean zero and variance $\sigma^2 t$, and changes in successive time intervals are independent.

What does Brownian Motion look like?



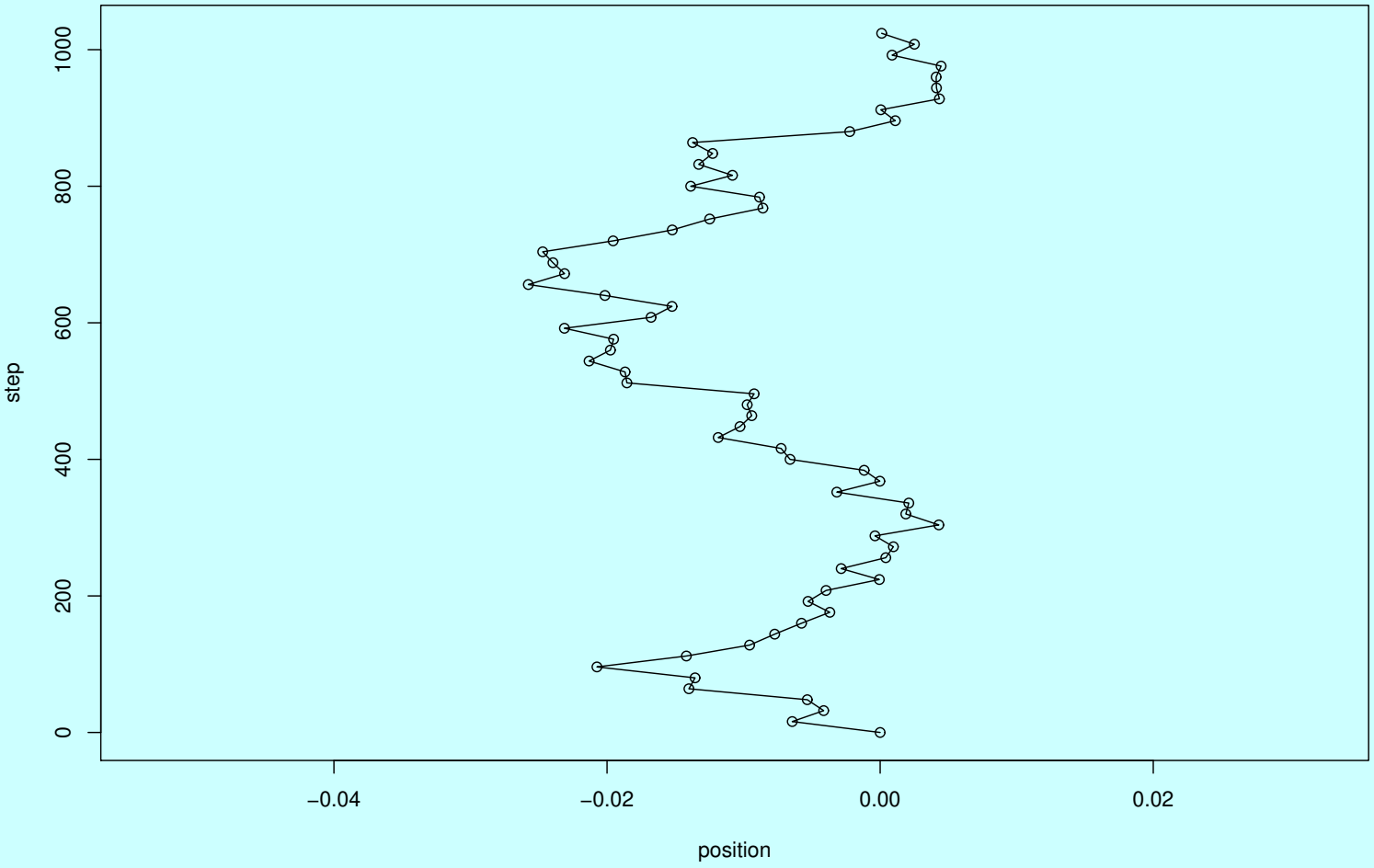
A random walk of 16 normally-distributed steps

What does Brownian Motion look like?



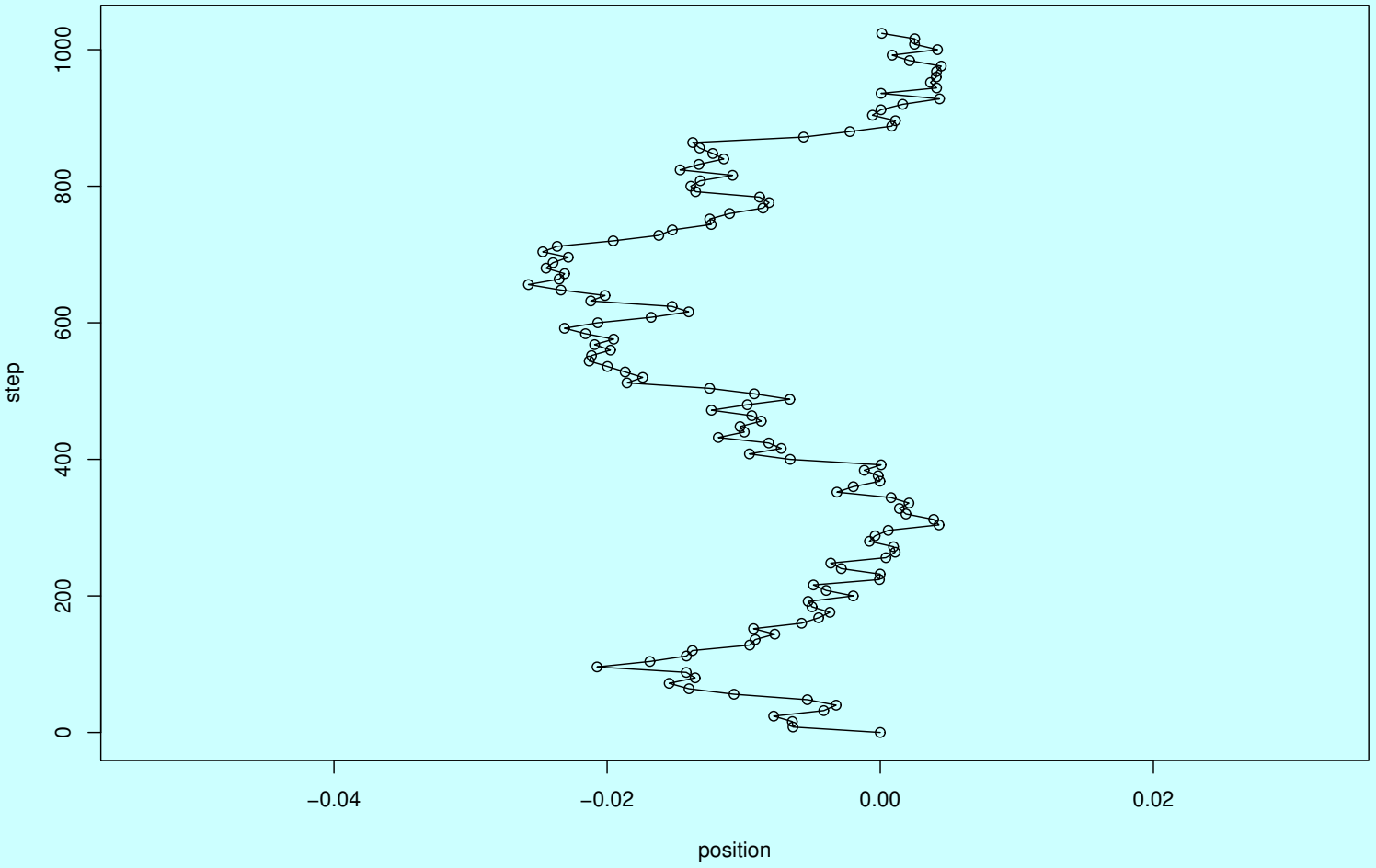
A random walk of 32 normally-distributed steps

What does Brownian Motion look like?



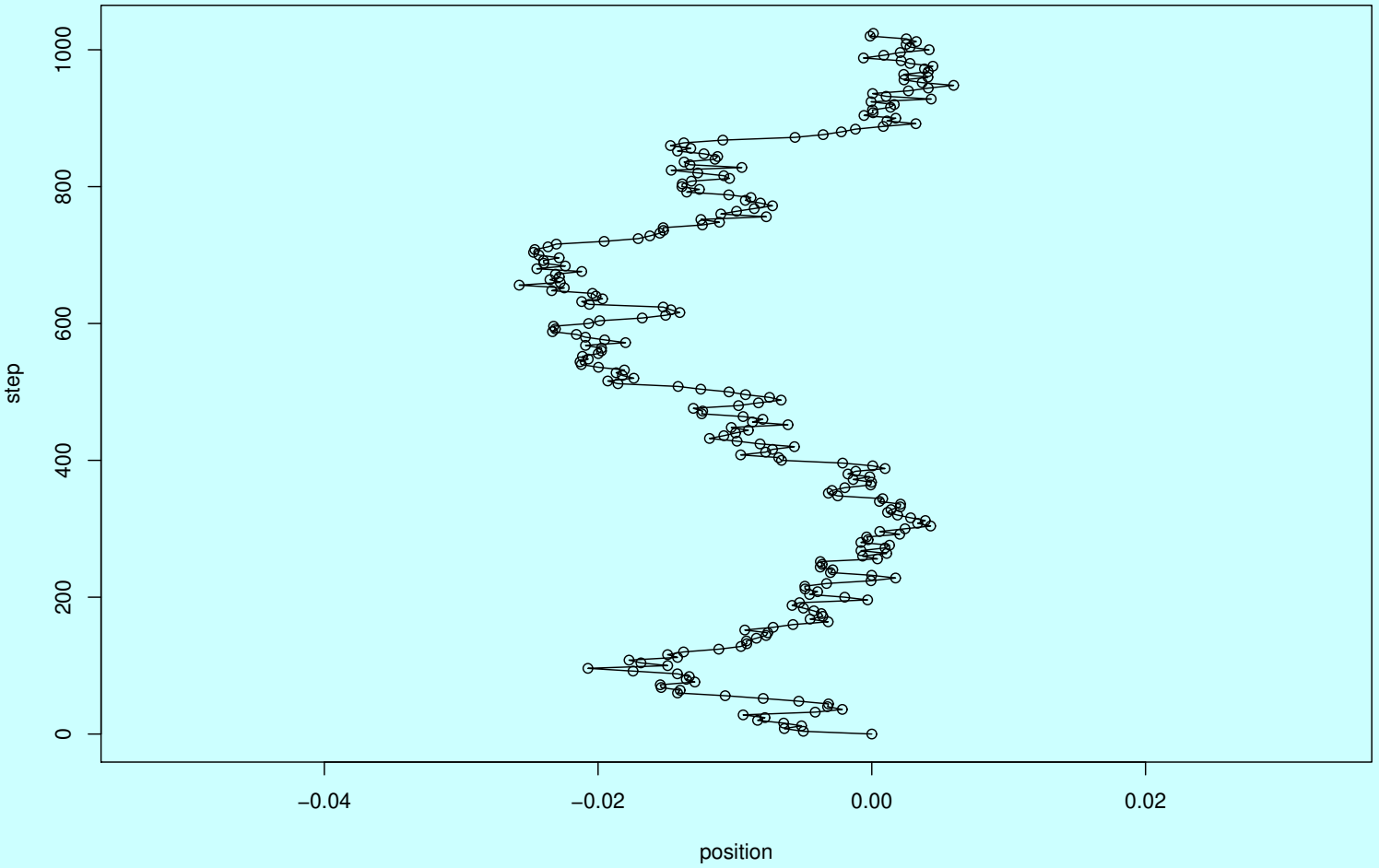
A random walk of 64 normally-distributed steps

What does Brownian Motion look like?



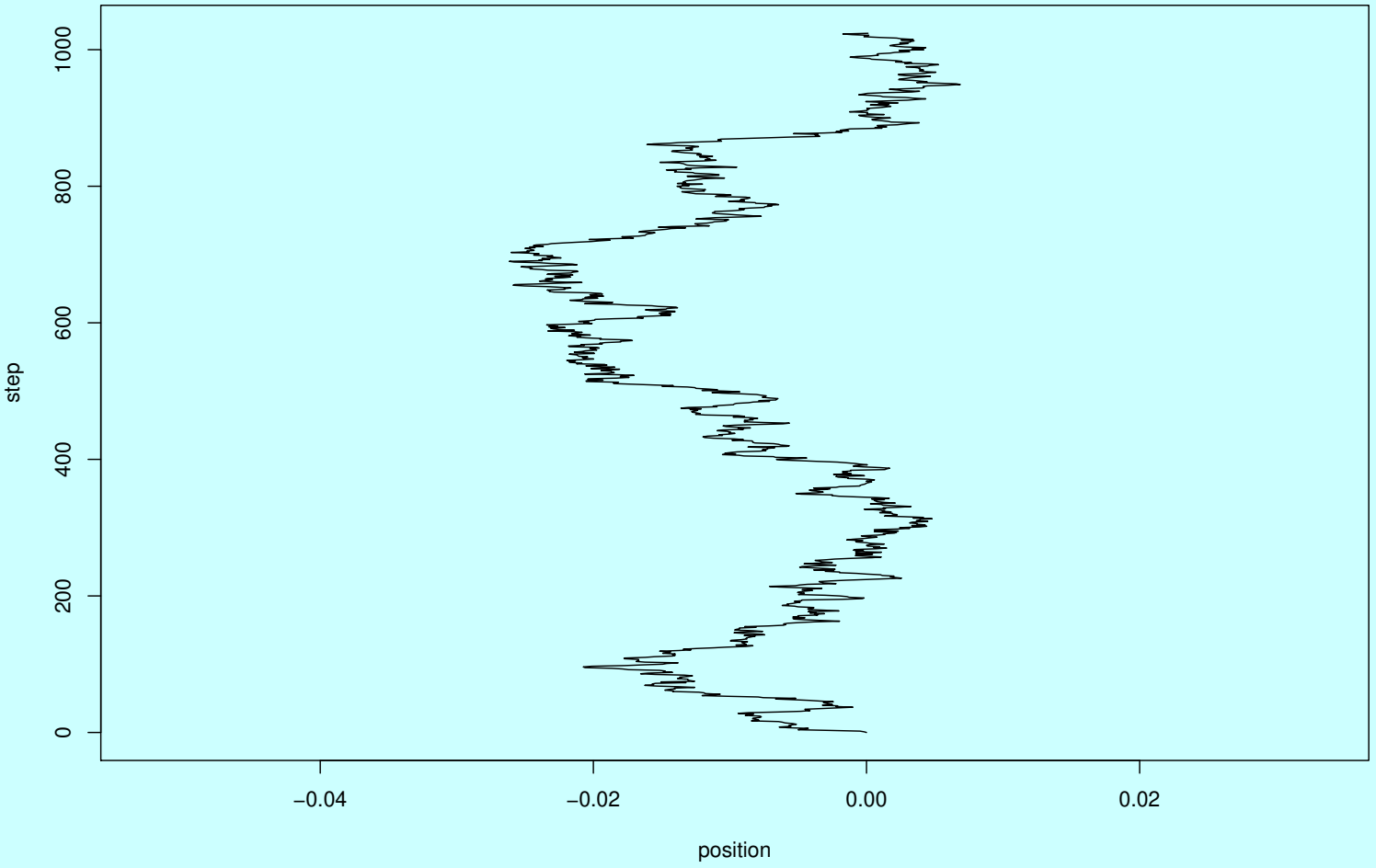
A random walk of 128 normally-distributed steps

What does Brownian Motion look like?



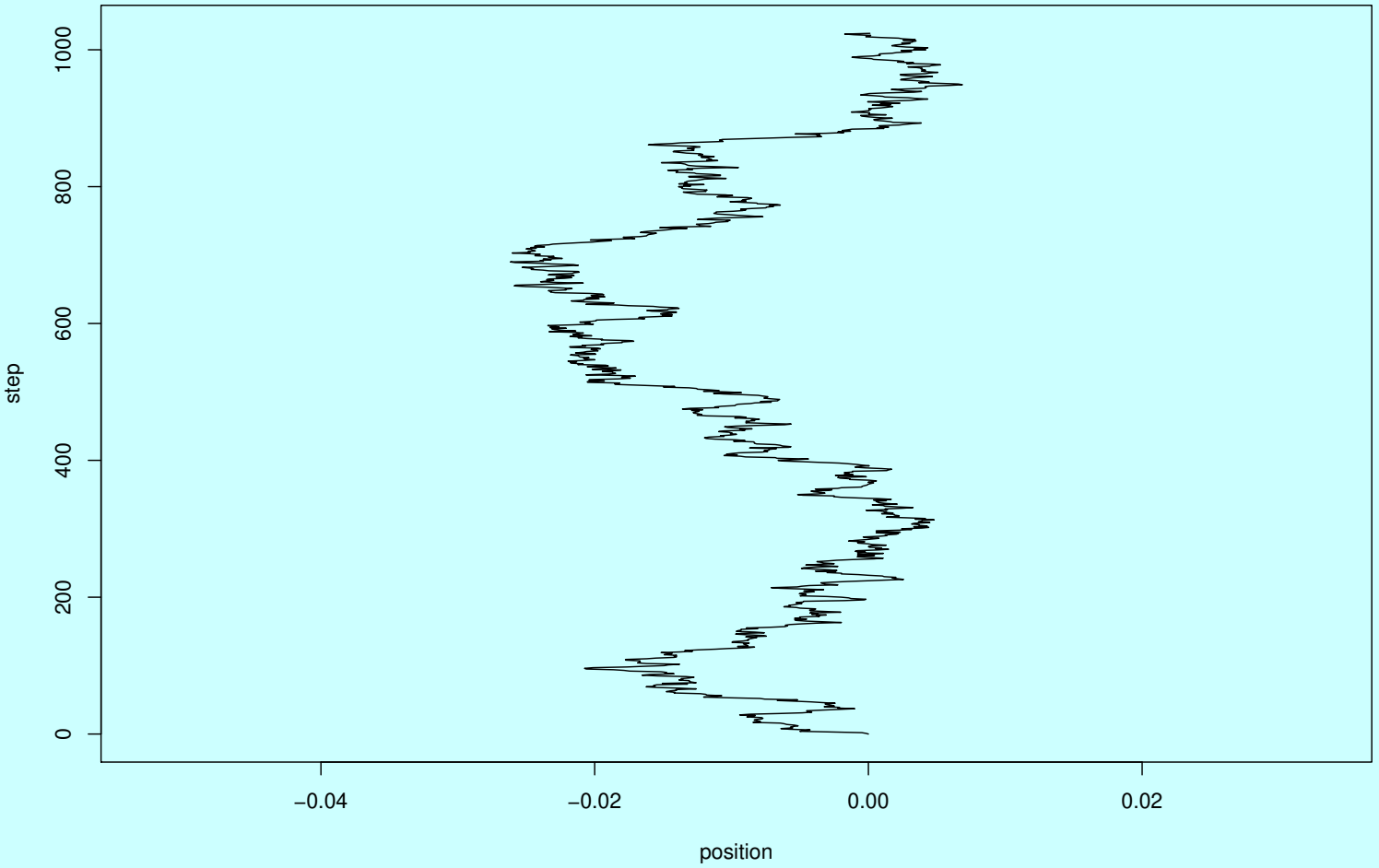
A random walk of 256 normally-distributed steps

What does Brownian Motion look like?



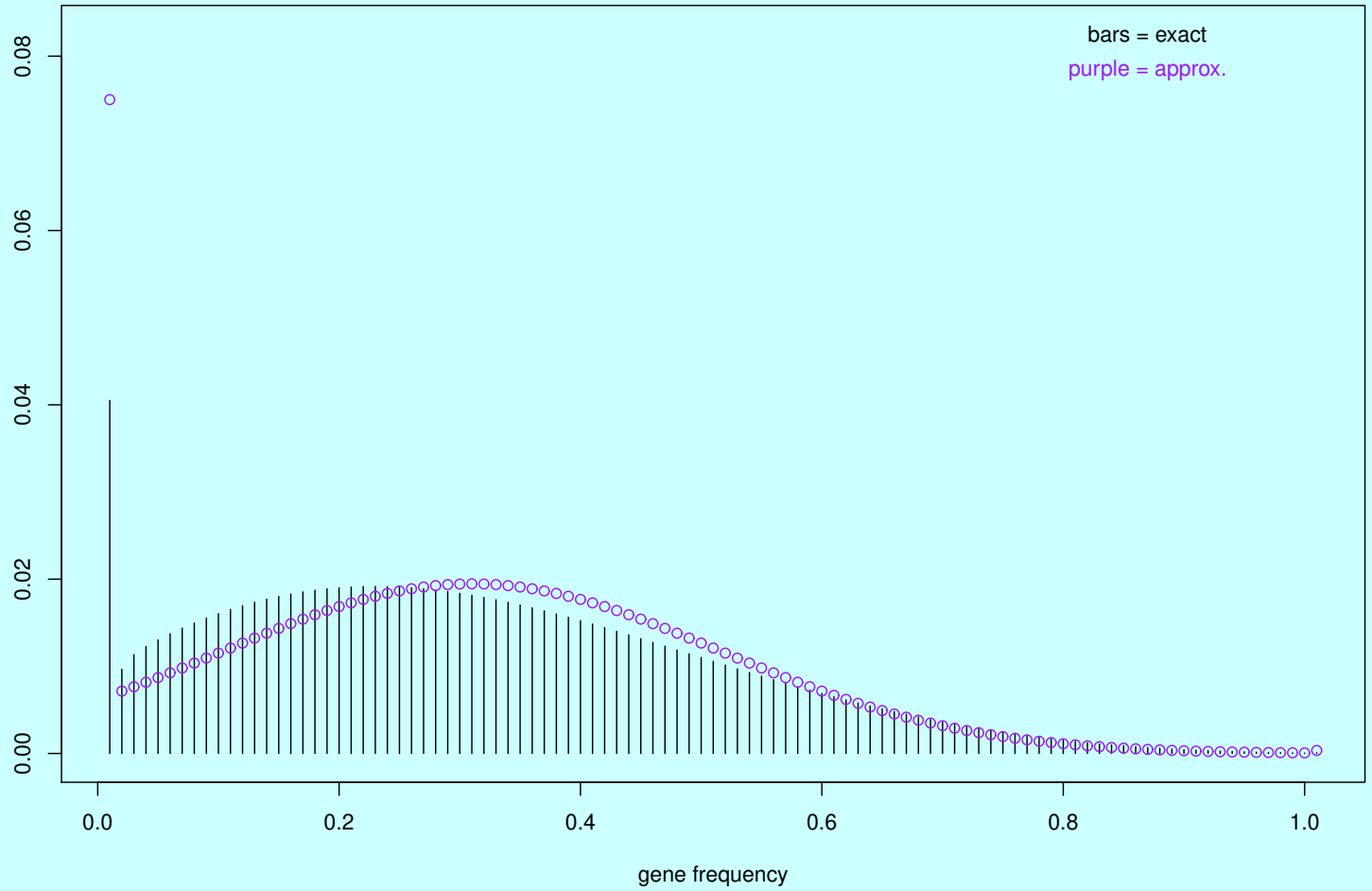
A random walk of 512 normally-distributed steps

What does Brownian Motion look like?



A random walk of 1024 normally-distributed steps

How good is this?



Starting with $p = 0.3$, after 20 generations in a population of size 50 (in a Wright-Fisher model). Compared with Brownian Motion with the same variance as at 0.3, binned.

What about a quantitative character?

If a quantitative character is a sum of contributions from a number of loci, then if the individual locus gene frequencies have their change approximated by Brownian Motion, the linear combination will also change by Brownian motion. This works for multiple alleles.

- if there is any dominance, there will be some nonlinearity and the approximation will be less good.
- Epistasis can cause even more trouble.

First discussed by me (Felsenstein, 1973).

Covarying character change along a lineage

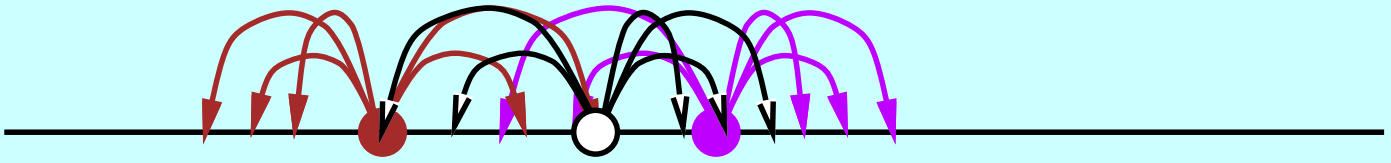
What is the distribution of changes in multiple characters (say p of them) along a lineage? Simply the appropriate multiple of the infinitesimal rate of change per unit branch length.

If a set of characters $\mathbf{x}$, changes under covarying Brownian motion, in time t (or a pseudo-time branch-length t) the change will be distributed as

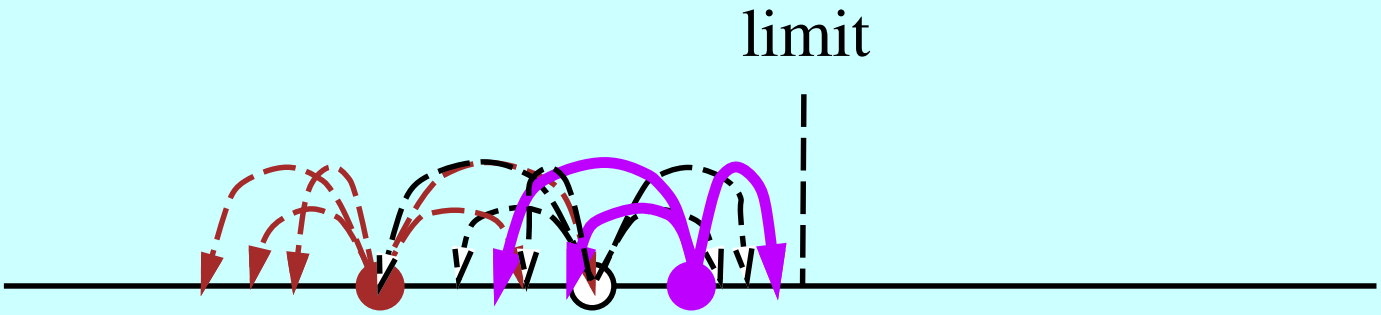
$$\mathbf{\Delta x} \sim \mathcal{N}(\mathbf{0}, \mathbf{V}t),$$

(where $\mathbf{V}$ is the covariance matrix of the infinitesimal change of the Brownian Motion).

But, if there are mutations making incremental changes ...



... as we saw with the discussion of quantitative characters, if a relatively constant genetic variance is maintained, and mutations have additive effects, then genetic drift will cause the mean to change in a random walk close to Brownian Motion.



However, if one approaches some limit where most mutations oppose movement to it, and there are no mutations allowing you to go past that limit, this approximation will be poor.

What causes change in quantitative characters?

For neutral mutation and genetic drift, can show that for a quantitative character with additive genetic variance V_A and population size N the genetic (additive) value of the population mean is the mean of the additive values of individuals. The variance of these means, from generation to generation, is then $1/N$ times the additive genetic variance:

$$\mathrm{Var}(\Delta\bar{g}) = V_A/N$$

The mean of the N additive values is expectation for the next generation. Genetic drift erodes the additive genetic variance by a fraction $1/(2N)$ each generation, and mutation pumps it up by a "mutational variance" V_M

If mutation and drift are at equilibrium:

$$E\left[V_A^{(t+1)}\right] = V_A^{(t)}\left(1 - \frac{1}{2N}\right) + V_M$$

In neutral traits additive genetic variance rules

so that, equating the V_A in the two generations, we can solve for the equilibrium genetic variance.

$$E[V_A] = 2NV_M$$

whereby

$$\mathrm{Var}[\Delta\bar{g}] = (2NV_M)/N = 2V_M$$

an analog of Kimura's result for neutral mutation.

Thus to transform characters to independent Brownian motions of equal evolutionary variance, we could use their additive genetic variance V_A.

With multiple characters ...

There is a precise analogue of this for multiple characters:

$$\mathrm{E}\left[\mathbf{A}^{(\mathrm{t}+1)}\right] \;=\; \mathbf{A}^{(\mathrm{t})}\left(1-\frac{1}{2\mathrm{N}}\right)+\mathbf{M}$$

where $\mathbf{A}$ is the additive genetic covariances, and $\mathbf{M}$ is the covariance matrix of pleiotropic effects of mutation.

$$\mathrm{E}\left[\mathbf{A}\right] \;=\; 2\mathrm{N}\,\mathbf{M}$$

and

$$\mathrm{Var}[\boldsymbol{\Delta}\bar{\mathbf{g}}] \;=\; \left(2\mathrm{N}\mathbf{M}\right)/\mathrm{N} \;=\; 2\mathbf{M}$$

so as long as mutations cause expected change zero (i.e. they are not near some biological limit), the effect of genetic drift is that the mean phenotype wanders according to the mutational covariances.

With selection ... life is harder

We've already seen, yesterday, that with natural selection the "Breeder's Equation" of Wright and Fisher (1920's)

$$\Delta z \;=\; h^2 S$$

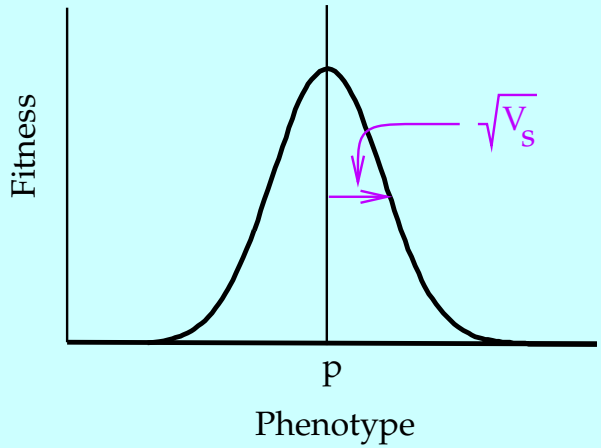
can be replaced by Russ Lande's (1976) recasting of that in terms of slopes of mean fitness surfaces (adaptive landscapes):

$$S \;=\; V_P \, \frac{d \log(\bar{w})}{d\bar{x}}$$

$$\Delta z \;=\; (V_A/V_P)\, V_P \, \frac{d \log(\bar{w})}{d\bar{x}} \;=\; V_A \, \frac{d \log(\bar{w})}{d\bar{x}}$$

Note – it's heritability times the slope of log of *mean* fitness with respect to *mean* phenotype. There is an exact multivariate analog of this equation.

Selection towards an optimum in one character



If fitness as a function of phenotype is:

$$w(x) = \exp\left[-\frac{(x-p)^2}{2V_s}\right]$$

Then to compute the effect of a fitness surface which is normal and multiplies the distribution, pointwise: after some completing of squares and integrating, the change of mean phenotype the mean phenotype m "chases" the optimum:

$$m' - m = \frac{V_A}{V_s + V_P}(p - m)$$

(There is an exact matrix analog of this for multiple characters).

Brownian motion from moving adaptive peaks

A reminder: we have seen earlier, and will hear about later: a process close to Brownian motion can also arise from natural selection toward an optimum which is itself moving. (But only if we can assume that the movement of the peak is itself Brownian Motion – a conclusion that would have to result from iour having a model of the environment itself).

Sources of evolutionary correlation among characters

Variation (and covariation) in change of characters occurs for two reasons:

- **Genetic covariances.** (the same loci affect two or more traits)

Sources of evolutionary correlation among characters

Variation (and covariation) in change of characters occurs for two reasons:

- **Genetic covariances**. (the same loci affect two or more traits)
- **Selective covariances** (Tedin, 1926; Stebbins 1950). The same environmental conditions select changes in two or more traits – even though they may have no genetic covariance.

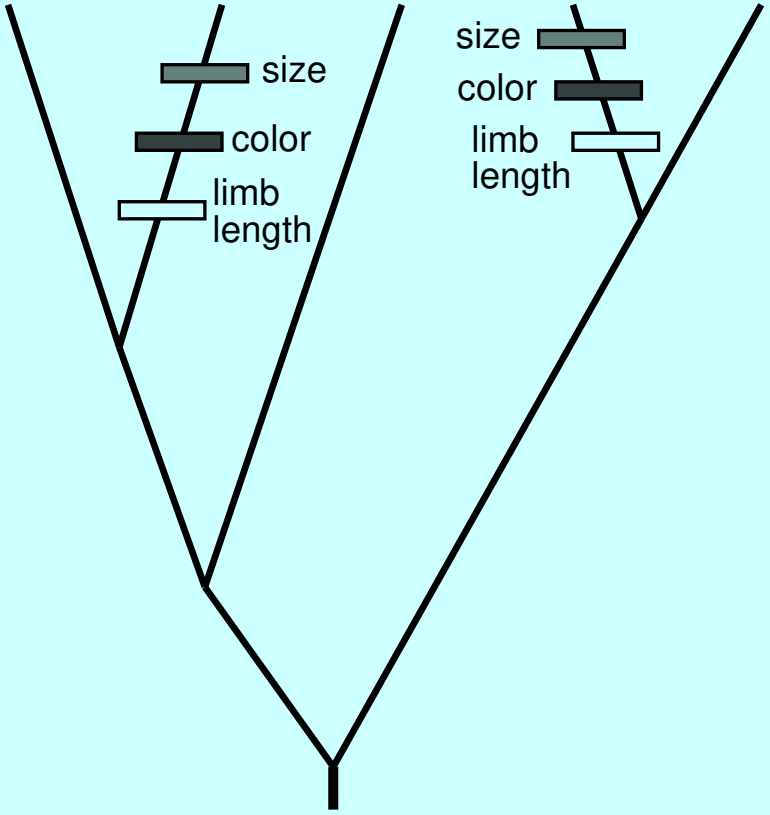
A simple example of selective covariance

covariation due not to genetic correlation
but to covariation of the selection pressure

These are Bergmann's, Allen's and Glogler's Rules
They are presumably not the result of genetic correlations
but result from patterns of selection

a simple example:

(temperate) (arctic) (temperate) (arctic) (temperate)



G. L. Stebbins. 1950. *Variation and evolution in plants.* Columbia Univ. Press, New York. page 121

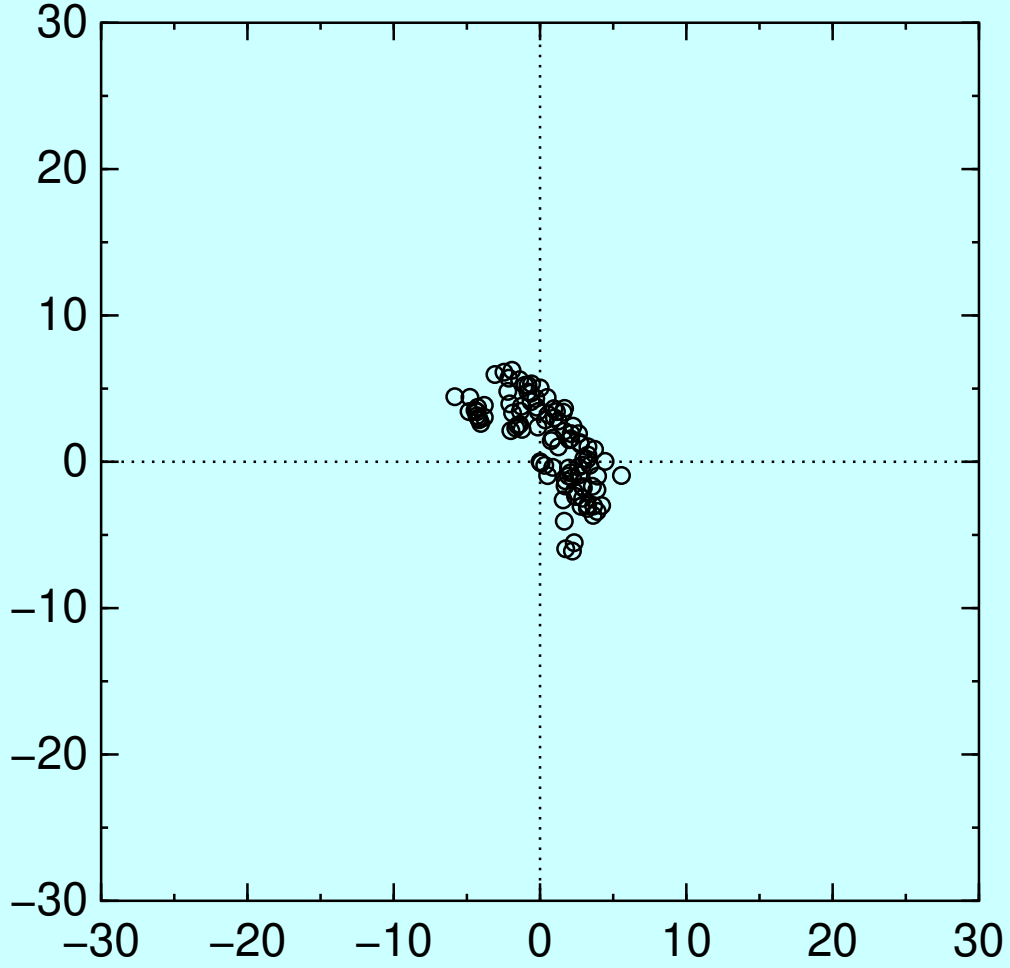
(thanks to Scott Armbruster for references)

Correcting for correlations among characters

Can we transform the set of characters to remove their correlations and thus end up with independent Brownian motions of equal variance?

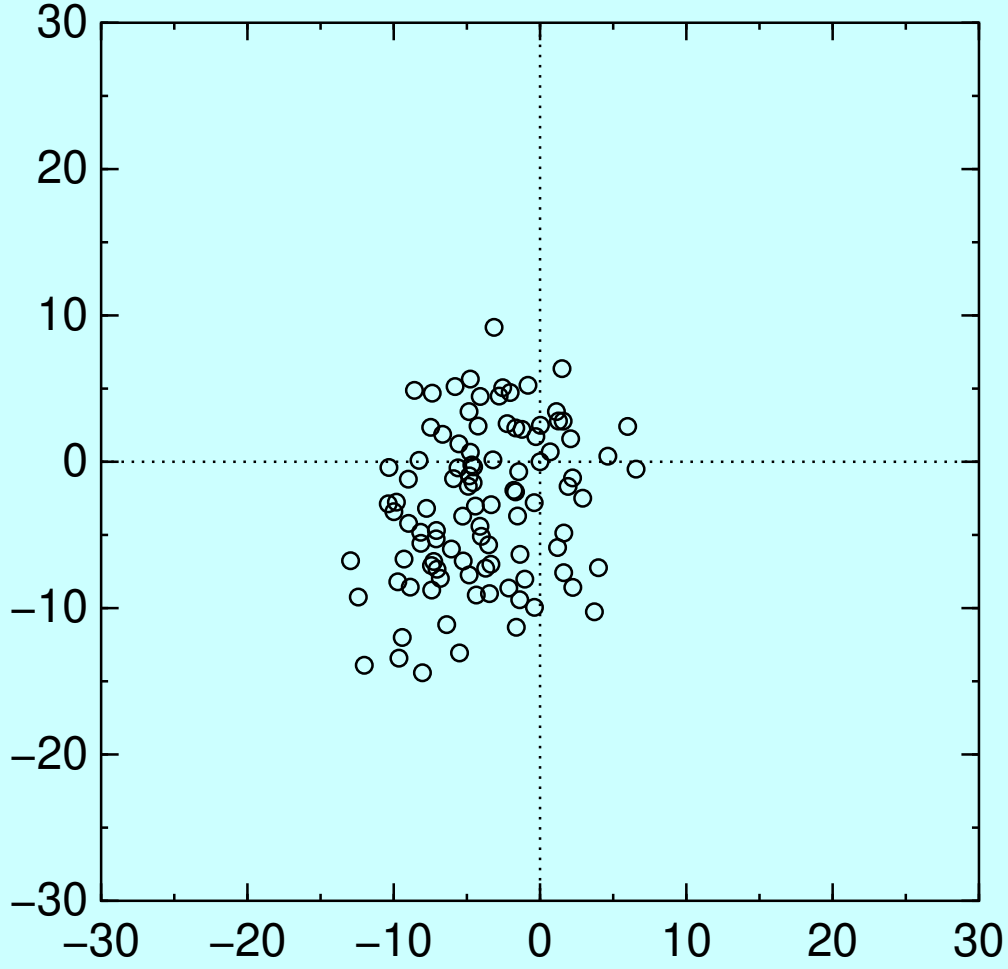
- We might hope to infer additive genetic covariances by doing quantitive genetics breeding experiments to infer them from covariances among relatives.
- There is little or no hope of inferring “selective correlations” without a complete understanding of the functional ecology.

A simulated example with two characters



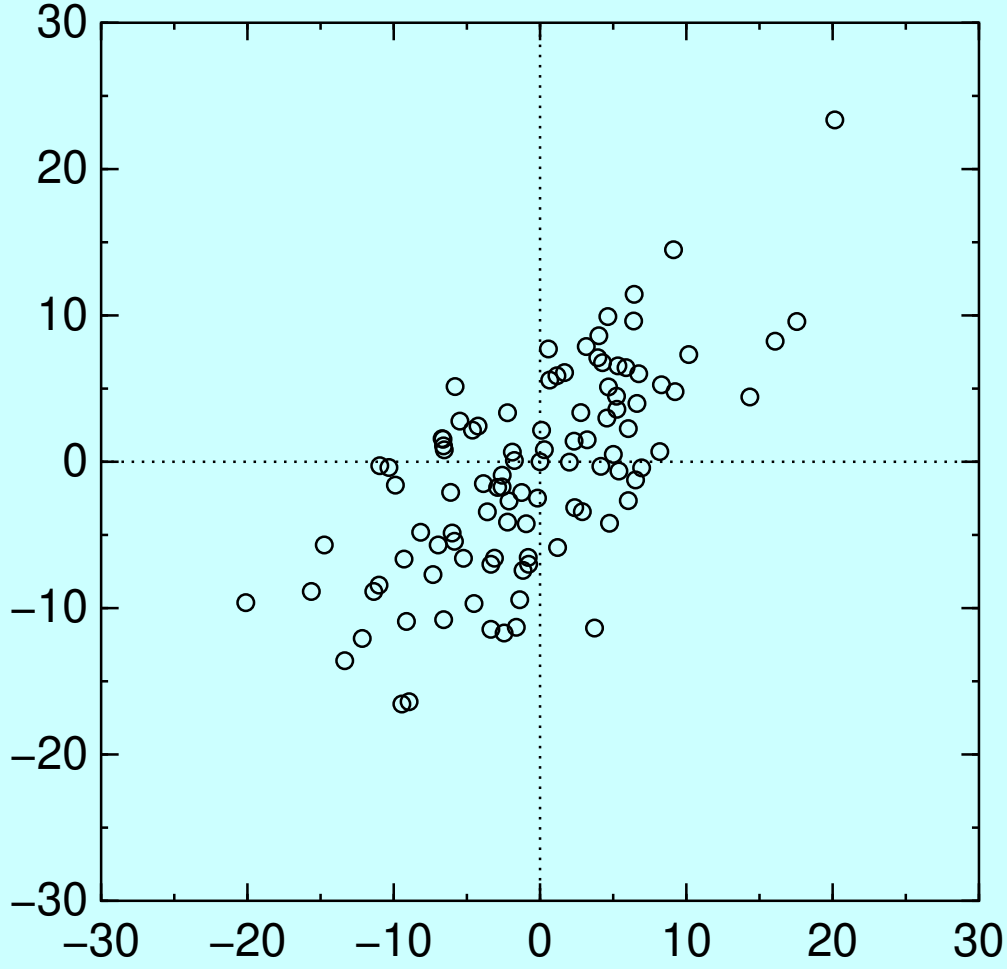
Genetic covariances are negative, but the wanderings of the adaptive peaks are positively correlated. In the first 100 generations the genetic covariances are most influential.

A simulated example with two characters



Genetic covariances are negative, but the wanderings of the adaptive peaks are positively correlated. After a while (every 10th generation up to generation 1000), the wanderings of the peaks start to be influential.

A simulated example with two characters



Genetic covariances are negative, but the wanderings of the adaptive peaks are positively correlated. In the long run (every 100th generation up to generation 10,000) the means go mostly where the peaks go.

A research program?

What we could imagine doing is:

- Infer genetic covariances from a quantitative genetics breeding experiment, perhaps in more than one species.
- Infer the covariances of the changes along the phylogeny.
- From them, back-calculate the selective covariances.

(For the case of chasing peaks, there is also hope of using differences among neighboring species to infer the genetic covariances).

References for the Brownian Motion approximation

Edwards, A.W. F. and L. L. Cavalli-Sforza. 1964. Reconstruction of evolutionary trees. pp. 67–76 in *Phenetic and Phylogenetic Classifcation*, ed. V. H. Heywood and J. McNeill. Systematics Association Publ. No. 6, London.' **[The first paper on numerical approaches to phylogeny reconstruction; uses parsimony and proposes likelihood for gene frequency trees]**

Edwards, A.W. F. 1970. Estimation of the branch points of a branching diffusion process. *Journal of the Royal Statistical Society B* **32:** 155–174. **[More detailed consideration of the statistical properties of a maximum likelihood approach to gene frequency phylogenies]**

Felsenstein, J. 1973. Maximum likelihood estimation of evolutionary trees from continuous characters. *American Journal of Human Genetics* **25:** 471–492. **[REML approach to gene frequency phylogenies, including the contrasts algorithm for rapid computation of likelihood]**

References for multivariate Brownian motion

Felsenstein, J. 1988. Phylogenies and quantitative characters. *Annual Review of Ecology and Systematics* **19:** 445-471. **[Review with mention of Ornstein-Uhlenbeck model]**

Felsenstein, J. 2004. *Inferring Phylogenies*. Sinauer Associates, Sunderland, Massachusetts. **[Mentions this model and also sample size issues in contrasts method]**.

Lande, R. 1976. Natural selection and random genetic drift in phenotypic evolution. *Evolution* **30:** 314-334. **[Lande's classic paper on drift versus optimum selection]**

Lande, R. 1979. The quantitative genetic analysis of multivariate evolution, applied to brain-body size allometry. *Evolution* **33:** 402-416.

References

Lande, R. 1980. The genetic covariance between characters maintained by pleiotropic mutations. *Genetics* **94:** 203-215.

Lynch, M. and W. G. Hill. 1986. Phenotypic evolution by neutral mutation. *Evolution* **40:** 915-935.

Stebbins, G. L. 1950. *Variation and Evolution in Plants*. Columbia University Press, New York. **[Describes selective covariance and cites Tedin (1925) for it]**

Tedin, 0. 1925. Vererbung, Variation, und Systematik der Gattung *Camelina*. *Hereditas* **6:** 275-386.

Armbruster, W. S. 1996. Causes of covariation of phenotypic traits among populations. *Journal of Evolutionary Biology* **9:** 261-276. **[Good exposition of selective covariance. He introduced me to that term.]**