

Provided for non-commercial research and education use.  
Not for reproduction, distribution or commercial use.



This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/copyright>



# Natural selection of altruism in inelastic viscous homogeneous populations

Alan Grafen\*, Marco Archetti

*Zoology Department, Oxford OX1 3PS, UK*

Received 9 October 2007; received in revised form 22 January 2008; accepted 22 January 2008

Available online 31 January 2008

## Abstract

Biological explanations are given of three main uninterpreted theoretical results on the selection of altruism in inelastic viscous homogeneous populations, namely that non-overlapping generations hinder the evolution of altruism, fecundity effects are more conducive to altruism than survival effects, and one demographic regime (so-called death–birth) permits altruism whereas another (so-called birth–death) does not. The central idea is ‘circles of compensation’, which measure how far the effects of density dependence extend from a focal individual. Relatednesses can then be calculated that compensate for density dependence. There is very generally a ‘balancing circle of compensation’, at which the viscosity of the population slows up selection of altruism, but does not affect its direction, and this holds for altruism towards any individual, not just immediate neighbours. These explanations are possible because of recent advances in the theory of inclusive fitness on graphs. The assumption of node bitransitivity in that recent theory is relaxed to node transitivity and symmetry of the dispersal matrix, and new formulae show how to calculate relatedness from dispersal and vice versa. © 2008 Elsevier Ltd. All rights reserved.

*Keywords:* Inclusive fitness; Population viscosity; Dispersal; Relatedness; Density dependence

## 1. Introduction

Recent work of Taylor et al. (2007a, henceforth TDW) has uncovered a striking natural structure in the theory of the natural selection of altruism in inelastic homogeneous populations, that is where the total population size is fixed and all individuals are equivalent. Here, we develop this natural structure, creating concepts and establishing mathematical results, focussing on calculating the central concept of relatedness, and basing interpretations on it. The new ideas prove their worth by allowing an explanation for each of three main existing results about selection in viscous populations, namely that the Fisher–Wright demographic process is not conducive to neighbourly altruism while the Moran demographic process is (Taylor and Irwin, 2000; Irwin and Taylor, 2001); that fecundity effects are more conducive to altruism than survival effects;

and that within the Moran process, the ‘birth–death’ replacement scheme is not conducive to neighbourly altruism while the ‘death–birth’ scheme is (Ohtsuki et al., 2006; Taylor et al., 2007a). There are more fundamental results about ‘Isolation by Distance’ models, recently reviewed by Rousset (2004), which underpin and justify many of the arguments in this paper, but which do not require any further explanation. The three mentioned have had no clear biological interpretation, and are the main unexplained results in the literature: a single explanation is now offered here.

The balance between relatedness to neighbours and increased competition between neighbours has been much discussed (e.g. Hamilton, 1971; Grafen, 1984; Queller, 1994; Kelly, 1994; Rousset, 2004; Gardner and West, 2004). Earlier work tended to show that the two forces cancelled each other out (Taylor, 1992a, b; Wilson et al., 1992). Much recent work has focussed on more complex situations, with elastic environments (Lehmann et al., 2006) sometimes associated with parasitism (Gardner et al., 2004) or non-independent dispersal through budding (Gardner and West, 2006), multiple traits such as punishment

\*Corresponding author at: St. John's College, Oxford OX1 3JP, UK. Tel.: +44 1865 277438; fax: +44 1865 277435.

E-mail addresses: [alan.grafen@sjc.ox.ac.uk](mailto:alan.grafen@sjc.ox.ac.uk) (A. Grafen), [marco.archetti@zoo.ox.ac.uk](mailto:marco.archetti@zoo.ox.ac.uk) (M. Archetti).

(Lehmann et al., 2007b), or finite but non-Poisson offspring number (Lehmann and Balloux, 2007). The strong results of TDW, and as developed here, are in the inelastic homogeneous case, but the simple theory is one important preparation for the complex.

An important technical result is the relaxation of TDW's condition for what counts as homogeneity of a population. They required node bitransitivity, to ensure that the dispersal matrix and the identity-by-descent matrix commute. Here we show that node transitivity combined with symmetry of the dispersal matrix will do instead. Indeed, the condition actually required is strictly weaker still, but less biologically meaningful. This technical work is conducted in Appendix B. It relies on being able to measure relatedness in stationary populations (see Rousset, 2004, p. 120), and its interpretation is greatly enhanced by the knowing that an expected increase in gene frequency, an enhanced probability of fixation and a higher inclusive fitness are all equivalent under weak selection (Rousset and Billiard, 2000; Taylor et al., 2000; Rousset, 2004; Wild and Taylor, 2004), which we assume throughout.

In Section 2, we introduce the concepts of '*g*-circles' and compensated relatednesses, which are biologically meaningful ways of understanding the effects of density dependence on the selection of altruism. Section 3 looks at the properties of compensated relatednesses under the Moran and Fisher–Wright demographic schemes, while Section 4 provides numerical illustrations, considers whether it is useful to discuss altruism in terms of the original parameters *b* and *c* instead of the offspring number after density dependence, and explains the difference between the death–birth and birth–death replacement schemes. The discussion in Section 5 considers the limitations of the theory. Appendix A finds explicit equations for uncompensated and compensated relatedness in terms of dispersal and vice versa, providing a firm theoretical understanding of the model.

It is worth noting here that we use relatedness based on identity by state, following Hamilton (1970) and many later authors, which measures a statistical association between alleles; rather than relatedness based on pedigrees as first used in connection with social behaviour by Hamilton (1964).

## 2. *g*-circles and compensated relatedness

We begin by reviewing how graph theory is applied to population modelling by Taylor et al. (2007a). The simplest interpretation, and the one we shall follow, is that the range of the species is divided into territories or areas that are occupied by one individual at a time. In graph theory, each territory is represented by a *node*. When an individual at node *i* dies, she is replaced, and the parent of the replacement comes from different nodes with different probabilities. The notation  $d_{ij}$  means the probability that the parent of the individual at node *i* occupied node *j* (TDW insisted that an individual could not replace herself,

so  $d_{ii} = 0$ , but we do not make that requirement here). As an extension, we also use  $d_{ij}^{(2)}$  for the probability that the grandparent of the individual at node *i* occupied node *j*, and in general  $d_{ij}^{(g)}$  for the probability that the *g*th ancestor did. These probabilities apply in the null model without selection, where all individuals are playing the same strategy. We assume that  $d_{ij}$  satisfy two simple symmetry assumptions—(i) that  $d_{ij} = d_{ji}$  and (ii) node transitivity, which says all nodes look the same. (More precisely, node transitivity means that for each pair of nodes, there exists a permutation of all nodes that carries the first of the pair to the second, and leaves the dispersal matrix unchanged.) For a formal statement, and a discussion of the weakening compared to TDW's assumptions, see Appendix B.

The basic circumstance of inclusive fitness is that an actor accepts a cost *c* and donates a benefit *b* to a recipient. But in a population of constant size, the benefit *b* must be compensated somewhere by losses to other individuals. Sometimes authors speak of a 'bland density dependence' in which all individuals in the population share equally in these compensating losses. But density dependence is likely to be local in a viscous population, and the crucial question is: who are those individuals that pay the compensation? There is no fixed answer, but we construct a series of 'circles of neighbours', called the 0-circle, 1-circle, 2-circle, etc., and assume that the compensation will be paid by 1-circle, or possibly spread across the circles. Circle is not intended in a geometric sense, but as in the ordinary language phrase 'social circle'. Each circle is like a probability distribution over the population, and individuals can belong to quantitatively different extents. The general principle is that the *g*-circle of the recipient contains nodes according to the probability that they contained the *g*th ancestor of the recipient. Under our symmetry assumptions, this is equal to the probability that the node's *g*th ancestor occupied the recipient's node. Thus, the 0-circle is just the recipient herself. The 1-circle of the recipient contains individuals according to the probability that, should the recipient die, that individual would supply the offspring that replaced her. We can think of those individuals as the likely parents (or offspring) of the recipient or, more accurately, as occupying the nodes that were occupied by the parent (or offspring) of the recipient. The 2-circles are the likely grandparents (or grandoffspring).

How might circles be relevant to understanding density-dependent compensation? If the cost to an individual is a chance of dying, then the beneficiaries will be the 1-circle, who will supply the offspring to replace the individual. Or, suppose an individual produces more offspring, who disperse as usual. If they compete with existing adults in the 1-circle, then the success of the offspring is the loss of those members of the 1-circle. If instead they compete with offspring who have dispersed from *their* parents' nodes, compensation falls on the 2-circle. This is because the aggregate of the 1-circle around the members of an individual's 1-circle is precisely the 2-circle of that original

individual. Moving outwards, if one individual could help its offspring to have more offspring themselves (for example, by extra parental care or tuition in parenting skills), then compensation could spread to the 3-circle. If offspring of different parents compete, and move away from more to less competitive arenas, then competition could be spread to more distant circles. Thus the circles are flexible constructs for describing the impact of density-dependent compensation.

Applying the concept of  $g$ -circle to an island model may help explain it further. Suppose we have  $N$  individuals in each of  $M$  islands, and that an individual is replaced by an offspring of a random fellow-islander with probability  $\beta$  and of a random member of the population with probability  $1 - \beta$ . In each case, the random choice includes the individual itself. Then the 1-circle contains self and each fellow-islander with a weight of  $\beta/N + (1 - \beta)/(MN)$ , and each other individual with a weight of  $(1 - \beta)/(MN)$ . The  $g$ -circle alters those weights to  $\beta^g/N + (1 - \beta^g)/(MN)$  and  $(1 - \beta^g)/(MN)$ , respectively. Provided  $\beta < 1$ , therefore, the weighting shifts to diminish the difference between fellow-islanders and others.

The second concept is ‘compensated relatedness’, short for ‘density-dependent compensated relatedness’. Suppose an actor gives  $b$  to a recipient with relatedness  $r$ , and that the density-dependent compensation is paid by some circle of the recipient, whose average relatedness to the actor is  $r_e$ . Then the two effects are  $rb$  and  $-r_e b$ . The net effect for the inclusive fitness of the actor is  $(r - r_e)b$ , and it is natural to call  $r - r_e$  the compensated relatedness of the actor to the recipient. Of course,  $r_e$  will depend on which  $g$ -circle pays the compensation, but there is no way of avoiding this complication. We will use the notation  $\tilde{r}_{ij(g)}$  to mean the compensated relatedness of individual  $i$  to individual  $j$ , when the fitness effect on  $j$  is compensated in her  $g$ -circle. The basic idea of subtracting the relatedness of those individuals that suffer density-dependent compensation from the relatedness of the primary individual to produce a net relatedness has been employed before (e.g. Taylor and Irwin, 2000; Gardner et al., 2004; Gardner and West, 2004, 2006), but the indexing by circle of compensation is new.

A more general notation is needed when the compensation is felt by circles according to a probability distribution  $\tau$ . For example,  $\tau_1 = \tau_2 = 1/2$  would mean that half the compensation was felt by the 1-circle and half by the 2-circle. We will refer to the  $\tau$ -circle as the mixture of the  $g$ -circles and write the compensated relatednesses as  $\tilde{r}_{ij(\tau)}$ . If at each generation offspring are distributed to circles according to the probability distribution  $\tau$ , then  $k$  generations later, there will be a particular distribution of the descendants. Mathematically, that distribution is the ‘ $k$ th self-convolution’ of  $\tau$ , and that concept will be needed below. Formally we can think of an integer value for  $\tau$  as representing the probability distribution over the integers which places all probability on that particular integer.

Similarly, the cost  $c$  paid by the actor will also have density-dependent effects, and so  $\tilde{r}_{ii(h)}$  will be the compen-

sated relatedness of the actor to herself when the density-dependent compensation will be paid by her own  $h$ -circle. We need to use a different symbol, because there is no necessary reason why the cost and benefit should be compensated by the same circle. For example, if an actor risks her life to help a neighbour reproduce, then it is likely that  $h = 1$  because her death would benefit the 1-circle, while the neighbour’s reproduction might be compensated by the 2-circle.

The compensated relatednesses as defined are useful, but lack the appealing property that relatedness to self is 1. It is therefore convenient to define a further system of relatednesses closely based on the first, but restoring that property. Using  $\tilde{\rho}$  for actor-referenced compensated relatednesses, we write

$$\tilde{\rho}_{ij(\sigma\tau)} = \frac{\tilde{r}_{ij(\tau)}}{\tilde{r}_{ii(\sigma)}}.$$

These relatednesses have the property that relatedness to self is always 1, provided compensation for actor and recipient is in the same circle, but they are in some ways more complicated. They are doubly-indexed by compensation circle, as it is necessary to state both the circle for the actor ( $\sigma$ ) and the circle for the recipient ( $\tau$ ). Some of the tables in Section 4 are in terms of  $\tilde{\rho}$  as this is sometimes easier to interpret.

We are now in a position to write down the inclusive fitness effect of an actor  $i$  in the classic position of accepting a cost  $c$  to give a benefit  $b$  to recipient  $j$ , where the density-dependent compensation will be paid by the actor’s  $h$ -circle and the recipient’s  $g$ -circle:

$$\tilde{r}_{ij(g)}b - \tilde{r}_{ii(h)}c$$

which has the same sign as

$$\tilde{\rho}_{ij(hg)}b - c. \tag{1}$$

Fundamental work (Rousset and Billiard, 2000; Taylor et al., 2000, 2007b; Rousset, 2004; Wild and Taylor, 2004), goes some way towards proving that selection proceeds according to this quantity in all the circumstances required, so we will conclude that the action will be selected for if the inclusive fitness effect is positive and against if it is negative. There are important subtleties in that literature, but irrelevant to present purposes, about exactly what being favoured by natural selection means in finite populations. Here we proceed in succeeding sections to study how  $\tilde{r}_{ij(g)}$  and  $\tilde{\rho}_{ij(hg)}$  behave, depending on the relative positions of individuals  $i$  and  $j$  and on the values of  $g$  and  $h$ .

### 3. Properties of compensated relatednesses

The main pattern in the balance between relatedness and competition is already clear in outline, in light of the  $g$ -circles introduced in the previous section. If compensation is distant enough, then the relatedness caused by viscosity will dominate, while if compensation is very local, competition will be a powerful force. But, how powerful?



And at what point does the balance tip between favouring altruism to neighbours and favouring spite towards neighbours? TDW provided the key to answering these questions, and here their answers are extended. After deriving general results, we look in more detail at the ‘Moran process’ in which one individual at a time dies and is replaced in the population, or the ‘Fisher–Wright process’, in which all individuals die and are replaced simultaneously (see Ewens, 2004, for origins and development of these processes).

The first generalisation is to adopt the Cannings process (Ewens, 2004) and allow a mixture of the Moran and Fisher–Wright processes (following e.g. Irwin and Taylor, 2001; Lehmann et al., 2007a). The Cannings process operates by supposing that at each ‘step’, each individual has an independent probability  $s$  of surviving, and  $1 - s$  of dying and being replaced. Replacement happens according to the dispersal matrix. We also introduce a notation  $c_{ij}$  that is defined as

$$c_{ij} = \frac{2s}{1+s} d_{ij} + \frac{1-s}{1+s} d_{ij}^{(2)} \quad (2)$$

which is interpreted as follows. Suppose we locate an ancestor of the individual at node  $i$ , to be more specific we choose to locate the parent with chance  $(2s)/(1+s)$  and the grandparent with chance  $(1-s)/(1+s)$ . Then the probability of that ancestor having occupied node  $j$  is  $c_{ij}$ .

It is now time to define the central concept of ‘balancing compensation’. This is a particular distribution of compensation between circles represented by  $\tau_1 = 2s/(1+s)$  and  $\tau_2 = (1-s)/(1+s)$ .  $\tau$  has up to now represented a general probability distribution, but from now on will refer to this particular one. Compensation in the  $\tau$ -circle is called balancing compensation because Eq. (10) of Appendix A generalises a result of TDW to show that the following compensated relatednesses hold in this case

$$(1 - r_\tau)^{-1} \tilde{r}_{ij(\tau)} = \begin{cases} 1 & i = j \\ \frac{-1}{n-1} & i \neq j, \end{cases} \quad (3)$$

where  $r_\tau$  is the relatedness of an individual to its own  $\tau$ -circle and  $n$  is the population size. The relatednesses that hold under balancing compensation have two remarkable features. The compensated relatedness to a neighbour is just the same as to any other individual in the population, so spatial structure of the population creates no spatial structure in the compensated relatednesses. Second, the values are in the same proportion (1 to self,  $-1/(n-1)$  to others) as in an unstructured, panmictic population. With  $\tau$ -circle compensation, therefore, the viscosity of the population should make no difference to the direction of natural selection of social behaviour, and this holds for altruism towards any individual, not just immediate neighbours. The speed of selection is slowed down by the factor  $1 - r_\tau$ , and this is not surprising. If all fitness effects are compensated locally, on related individuals, then the

net effect on an allele is inevitably proportionally reduced according to the strength of the relatedness.

This knife-edge result is also significant for what it suggests about either side of the blade. If compensation is closer to the actor than the balancing compensation, then neighbours are negatively valued, and closer neighbours more so. If compensation is more distant, then neighbours are positively valued, and closer neighbours more so.

The three major findings in the literature may be interpreted in terms of this balancing compensation. Taylor and Irwin (2000) and Irwin and Taylor (2001) showed that non-overlapping generations were less conducive to neighbourly altruism than overlapping generations, and that fitness effects through fecundity were more favourable to altruism than survival effects. These effects are explained as follows. Effects through survival have  $g = 1$  (i.e. density dependence falls on individuals one dispersal step away), while effects through fecundity have  $g = 2$  (i.e. density dependence falls on individuals two dispersal steps away). The balancing circle is the  $\tau$ -circle which is a mixture with  $2s/(1+s)$  of  $g = 1$  and  $(1-s)/(1+s)$  of  $g = 2$ . The density dependence of survival effects is always closer than that of fecundity effects, and so competition has a higher strength compared to relatedness. As generations overlap more, i.e.  $s$  increases, the balancing circle moves back towards the 1-circle, giving a larger margin of advantage to the fecundity effects at  $g = 2$ , and a smaller margin of disadvantage to the survival effects at  $g = 1$ . We can also see that neighbourly altruism is possible with non-overlapping generations provided the compensation is more widespread than  $g = 2$ . We will look more at this in Sections 3.2 and 4.

The third finding (Ohtsuki et al., 2006; Ohtsuki and Nowak, 2006; Grafen, 2007; Lehmann et al., 2007a) is that the so-called DB and BD demographic schemes differ in their capacity to support altruism. DB means that a random individual dies, and then neighbours compete to replace her, and their altruism affects those chances of replacement. BD means that an individual is chosen to reproduce according to the effects of altruism, and then a random neighbour dies to make space. As TDW point out, though not in these terms, compensation is in the 1-circle in the BD scheme but in the 2-circle in the DB scheme. Balancing compensation in the Moran scheme is in the 1-circle, and this implies altruism will not spread in the BD scheme, as  $g = 1$ , but will spread in the DB scheme, where  $g = 2$ . However, the order of birth and death cannot by itself be biologically significant, and we clear up this point when we look at numerical examples in Section 4.

The conclusions of this section hold very generally because the assumptions are very general. The population can be of any size, and the symmetry assumptions permit a wide range of structures. Roughly, cycles and grids in which immediate neighbours all have equal  $d_{ij}$  while all others are zero are included. But self-replacement is also allowed, as is a tiered structure in which near neighbours have a high chance of replacing the individual while more

distant neighbours have a lower chance. For a discussion of the limitations, see Section 5.

The rest of the paper looks at various special cases. The Moran and Fisher–Wright models have particular properties it is worthwhile to investigate in succeeding subsections. Then the following section looks at numerical examples.

### 3.1. The Moran process

Here we pursue further the patterns of compensated relatedness by restricting ourselves, for the moment, to the Moran process in which one individual at a time dies and is replaced. This corresponds to approaching the limit  $s = 1$  in the general case, so far as asymptotic relatednesses are concerned, but we must take a moment to establish this point. As  $s$  approaches 1, a higher and higher fraction of generations have zero deaths, but the population is completely unchanged after one of those generations. Of the generations with at least one death, a higher and higher fraction have exactly one death, and so the relatednesses increasingly approximate the Moran case. However, the Moran case is not the same as  $s = 1$  in all respects. The dynamics of probability of identity by descent are not the same, because the Moran case has exactly one death per generation while, as already mentioned, as the limit  $s = 1$  is approached, many generations have no deaths.

When we take the limit to  $s = 1$  in the general case, and when density-dependent compensation falls on the 1-circle, the relatedness and competition effects of viscosity exactly cancel each other out. TDW show that when density dependence is felt by the 2-circle, however, the cancellation is not complete. The values of the relatednesses  $\tilde{r}_{ij(2)}$  are instead as follows:

$$(1 - r_1)^{-1} \frac{n-1}{n} \tilde{r}_{ij(2)} = \begin{cases} 1 + d_{ii} - \frac{1}{n} & i = j \\ d_{ij} - \frac{1}{n} & i \neq j. \end{cases} \quad (4)$$

How should these values be interpreted? An individual values a neighbour according to the probability that they replace each other, minus  $1/n$ , which is the probability they would replace each other if replacement were completely at random. Thus an individual is valued according to its ‘excess replacement probability’. An individual values itself with an additional 1 over and above this excess replacement probability. Note also that relatedness to those with zero chance of replacing an individual is negative.

The directness of this result is interesting. Although relatedness builds up over generations, it seems that all the building up beyond the simple one-generation connection is cancelled out by the density-dependent compensation. Taking this further, the local nature of interactions means that overall gene frequencies are actually unimportant to the selection of neighbourly altruism. The key role played by the population gene frequency in the calculation of relatednesses (Grafen, 1985; Queller and Goodnight,

1989) has always seemed rather curious, as it suggests that if we study one population we cannot say how selection should proceed, but need to know the gene frequency of the whole species. The formulation of compensated relatednesses expresses a useful and affirming response to this concern about the relevance of far-flung populations. The discussion of this paragraph echoes that of Rousset (2004, p. 110).

The  $g = 3$  result provides  $\tilde{r}_{ij(3)}$  through

$$(1 - r_1)^{-1} \tilde{r}_{ij(3)} = \begin{cases} 1 + \left(d_{ii} - \frac{1}{n}\right) + \left(d_{ii}^{(2)} - \frac{1}{n}\right) & i = j \\ \left(d_{ij} - \frac{1}{n}\right) + \left(d_{ij}^{(2)} - \frac{1}{n}\right) & i \neq j. \end{cases}$$

Going out to a further circle has the effect of adding an extra term on the right-hand side. We add the probability of a double-dispersal-event link between the sites occupied by the individuals, again minus the probability expected in a completely random model.

It is clear that the pattern will be continued for higher  $g$  by adding in the ‘excess replacement probability in  $g$  steps’ to the right-hand side. It is interesting that individuals completely unconnected in  $g$  steps of dispersal acquire more and more negative relatednesses. Of course, the number of such individuals decreases as  $g$  increases, as eventually all individuals in a connected population will be kin of some kind, following the ‘inbreeding paradox’ of Seger (1981).

A final result in this section is about the ‘infinite circle’, i.e.  $g = \infty$ , where with a caveat about technical conditions of very limited empirical significance, the density dependence is compensated so remotely that it falls evenly on the whole population. Then, as shown in Appendix A, the compensated relatedness ( $\tilde{r}_{ij(\infty)}$ ) equals the uncompensated relatedness ( $r_{ij}$ ), as would be expected. This calculation provides an explicit formula for relatedness in terms of dispersal (and vice versa).

The main conclusion for the Moran process is that compensation in the 1-circle creates neighbourly neutrality, but that if compensation is spread any more widely then individuals should be altruistic to their neighbours, and more altruistic to closer neighbours. The wider the compensation is spread, the stronger that altruism should be, until with very wide compensation, neighbours should be treated according to their uncompensated relatednesses.

It was noted earlier that the parental probabilities and grandparental probabilities are more relevant, when compensation is felt locally, than the uncompensated relatednesses. Thus estimations of relatedness from genetic data cannot be automatically applied to the question of neighbourly altruism. One solution is to make separate field observations of the  $d_{ij}$ , to allow the compensated relatednesses to be calculated. An alternative is to use Eq. (11) to infer the  $d_{ij}$  from the relatednesses, obtaining the relatednesses from DNA samples, and so infer from molecular biological data what the pattern of neighbourly

altruism should be. There are many practical problems with this second possibility, but it seems a useful route, and could involve building on existing methods of calculating relatednesses (Queller and Goodnight, 1989) and Wright's *F*-statistics (Fontanillas et al., 2004) from genetic data. This possibility would be especially attractive if the framework could be extended in detail to diploid sexual organisms.

### 3.2. The Fisher–Wright process

Under the Fisher–Wright process, we take  $s = 0$  in Eqs. (2) and (3) to see that the balancing compensation is in the 2-circle. The compensation must therefore be spread more broadly under Fisher–Wright than under Moran to obtain neighbourly altruism. First we discuss two phenomena that occur in the Fisher–Wright process, and so increasingly as  $s$  approaches 0, namely inverted relatednesses to close neighbours, and the ‘sawtooth effect’.

Inverted relatednesses occur when compensation is to the 1-circle, for then competition outweighs relatedness, and local relatednesses are negative. Formally, it means that  $\tilde{r}_{ij(1)}$  will have negative values between close neighbours, and more negative than between distant neighbours. We look at this phenomenon in the numerical example in Section 4.

The ‘sawtooth effect’ is shown in Appendix C to be linked to the inverted relatednesses. If  $D$  is fully regular (see Appendix A), then the compensated relatednesses converge as  $g$  increases. But if  $D$  has an eigenvalue close to  $-1$ , then there are inverted relatednesses, and the convergence will be alternating. The sawtooth pattern seems likely to be rightly regarded as an artefact of the model, and unlikely to be encountered in nature. It is important to understand it when interpreting models, however. A linked effect is the ‘checkerboard effect’, discussed in the numerical examples in Section 4.

The simplest form of the checkerboard effect occurs on a torus with an even number of nodes in each of the basic cycles, which we imagine spreading out as a square, and think of as a chess board. If all dispersal is to one of the four immediate neighbour nodes, then all individuals on white squares leave offspring in black squares and vice versa, and therefore all the grandoffspring of individuals on white squares are on white squares. As shown in Appendix A, relatedness cannot even be defined in this case, at least not in the usual way. Partial effects arise in less extreme cases. For example, if the grid just mentioned has odd sides, then there is a distant connection between an individual and its immediate neighbours, and those relatednesses are actually the lowest, because the ancestral path must go all around the grid.

We now turn to the selection of neighbourly altruism in the Fisher–Wright process, and assume for the moment that compensation is felt in the 4-circle:

$$(1 - r_2)^{-1} \frac{n-1}{n} \tilde{r}_{ij(4)} = \begin{cases} 1 + d_{ii}^{(2)} - \frac{1}{n} & i = j \\ d_{ij}^{(2)} - \frac{1}{n} & i \neq j. \end{cases}$$

Thus  $g = 4$  has a reasonably similar pattern in the Fisher–Wright process to the  $g = 2$  case in the Moran process, with the substitution of grandparental probabilities for parental probabilities. Here we see the expected spatial pattern in which close neighbours are favoured more than distant neighbours. Fisher–Wright compensated relatednesses are likely to be lower for nearby individuals, as in general grandparental chances ( $d_{ij}^{(2)}$ ) are ‘smeared out’ spatially and so generally lower than parental chances ( $d_{ij}$ ).

The most important finding here is that the tendency found by Taylor and Irwin (2000) and Irwin and Taylor (2001) for non-overlapping generations to prevent neighbourly altruism is by no means absolute. There can be positive spite to neighbours at  $g = 1$ , there is indifference at  $g = 2$ , but neighbourly altruism can be selected, even if at levels likely to be somewhat lower than in the Moran case, when density-dependent compensation extends to  $g = 4$  or higher. Once compensation is very widely spread, then neighbours should be treated according to their uncompensated relatednesses, and these are generally lower than apply under the Moran process, but can still be substantial.

There is one factor that partially reverses some of this tendency for Fisher–Wright to produce less neighbourly altruism than Moran. So far, the comparisons have been made between the two processes at the same circle. Consider, however, the effect of an individual producing more offspring in the Fisher–Wright and Moran cases. In the Moran case, if those offspring compete with adults for their existing territories, then compensation is in the 1-circle. In the Fisher–Wright case this is impossible, and the offspring will be competing with other offspring from the 1-circle of the original individual's 1-circle, that is, compensation will fall on the 2-circle. Thus the very force that pushes out the balancing compensation to  $g = 2$  in the Fisher–Wright case, also means that at least in some cases, the fair comparison is between  $g = 1$  for Moran, and  $g = 2$  for Fisher–Wright. More generally, it may sometimes be fairer to compare the compensated relatednesses for the  $\tau$ -circle, and of course these are in the same proportions in the two schemes as well as spatially unstructured. None of this diminishes the significance of the lower uncompensated relatednesses to neighbours in the Fisher–Wright scheme, and this effect is bound to emerge in cases where compensation is spread very widely so that the compensated relatednesses approach the uncompensated relatednesses.

### 4. Numerical calculations on regular grids

The previous sections have been carried out in considerable generality. Here we apply the general formulae in simple grids and obtain numerical values for the various quantities. These grids are equivalent to the lattice models introduced by Malécot (1948, 1950) and reviewed by Rousset (2004, Chapter 3), who solves analytically for identity by descent coefficients using Fourier analysis. Here, we limit ourselves to showing numerical patterns of

relatedness, both compensated and uncompensated. In Section 4.1, the numerical examples are used to explain in a very concrete way what  $d_{ij}$  look like, and how compensated relatednesses are derived from relatednesses. Section 4.2 explains that, although the two demographic schemes (BD and DB) have different outcomes, this is, contrary to the suggestion of the names, not to do with the order of birth and death. Section 4.3 considers what is at stake when, as seems natural here, altruism is redefined in terms of the score rather than the fitness consequences.

We constructed cycles ( $k = 2$ ) and regular toroidal graphs with different numbers of neighbours ( $k = 3, 4, 6, 8$ ). The patterns of connections are illustrated in Fig. 1. Ohtsuki et al. (2006) set  $d_{ij} = 1/k$  for the  $k$  immediate neighbours. We follow them closely but not exactly in our choice of connection strengths. In order to avoid biologically unimportant problems of a technical nature, we permit a small probability ( $1/121$ ) that an individual will replace herself. The remaining probability  $120/121$  is divided equally between the  $k$  immediate neighbours. These patterns satisfy the assumptions of symmetry and node transitivity. The population size,  $n$ , is also varied, taking the values 64, 196 and 400. (Note that an explicit algebraic formula for relatednesses in the cycle with  $d_{ij} = 1/2$  between neighbours is given by Grafen, 2007.)

A number of other parameters were varied. The generational scheme was either Moran (one individual dies and is replaced at each step,  $s = 1$  in terms of the general model) or Fisher–Wright (all individuals die and are replaced at each step,  $s = 0$ ). We calculated relatednesses exactly in each case, by solving a matrix equation in the

rational, based on Eq. (8), using Mathematica 5.2 (Wolfram Research, Inc., 2005). Compensated relatednesses were then easy to derive, as shown in Appendix A.

The uncompensated relatednesses are of interest, not only as measuring how much genetic similarity to expect under different population structures, but also as the limiting case of compensated relatednesses as density-dependent compensation becomes very widely spread. Table 1 shows relatedness to the four closest type of neighbours across the range of structures studied. The greater the population size, the greater is  $r$  to neighbours, partly because the overall negative tendency of relatedness to another individual is lower ( $-1/(n-1)$ ), and partly because the neighbours are proportionally closer compared to the population as a whole. Relatedness to the most distant individuals becomes less negative as population size increases. Second, the more neighbours, the lower is  $r$ , as the genetic similarity is spread between more individuals. The Fisher–Wright results are affected by inverted relatednesses and the sawtooth effect, to be discussed shortly. Third, apart from the inverted relatednesses, relatedness decreases as neighbours become more distant, which is the primary geographical pattern.

Fig. 2 shows how uncompensated relatednesses vary over the whole population in special cases of the Moran and Fisher–Wright schemes, and also shows the difference between them. Nearby relatednesses are lower under Fisher–Wright, while the most distant relatednesses are less negative, preserving the important property that the sum of relatednesses over the population equals zero.

Compensated relatednesses to self are shown in Table 2. This shows the strength of the effect of the circle of compensation on the true cost of being altruistic, and it will be seen that rather rapidly the compensated relatedness approaches the uncompensated relatedness (which is 1). It also shows that we can have compensated relatednesses greater than 1, when the compensation is felt by a circle with a negative relatedness to the actor (Fisher–Wright cases with  $k = 3, 4$ ). The values in even generations go up, as the theory in Appendix C shows. The odd generations are affected by sawtoothing, as discussed in Section 3.2 and Appendix C. With larger  $k$ , neighbours of an individual are also each others' neighbours, introducing paths back to self with an odd number of segments, which diminishes the strength of this checkerboarding. Finally, as Rousset (2004, p. 4) points out, viscosity gives even apparently non-social behaviour a social aspect, as selection on non-social traits must still be compensated. The compensated relatednesses to self show how strong the selection is to choose the circle of compensation with the highest value of  $r$ . This will usually mean widening the circle, but while the sawtoothing is important, there are also sharp differences between adjacent even and odd circles, especially in the Fisher–Wright case and for small  $k$ .

Actor-referenced compensated relatednesses are shown in Table 3. We first point out that balancing compensation takes place, as shown analytically earlier, with  $g = 1$  for the

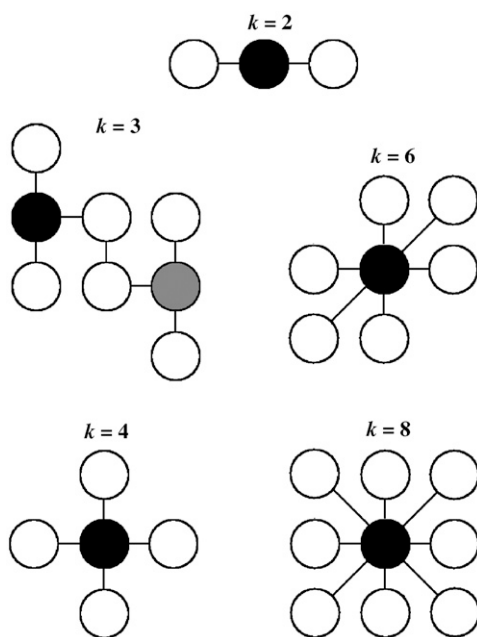


Fig. 1. The figure shows how an individual is connected to neighbours in a cycle ( $k = 2$ ), or in a square grid pattern ( $k \geq 3$ ) wrapped on a torus. With  $k = 3$ , there are two kinds of individuals, those connected to the neighbour on the right (shown in black), and those connected to the neighbour on the left (shown in grey).



Table 1

Relatednesses in some simple networks, with three different population sizes, to self ( $b = 1$ ), to the nearest neighbours ( $b = 2, 3$ ) and to the most distant individual in the population ( $b = b_{max}$ )

Population size	Circle	Model									
		Moran					Fisher–Wright				
		$k = 2$	$k = 3$	$k = 4$	$k = 6$	$k = 8$	$k = 2$	$k = 3$	$k = 4$	$k = 6$	$k = 8$
$n = 64$	$b = 1$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	$b = 2$	0.908	0.479	0.351	0.286	0.228	0.318	−0.182	−0.231	0.093	0.079
	$b = 3$	0.818	0.248	0.181	0.085	0.185	0.776	0.388	0.371	0.063	0.040
	$b = 4$	0.732	0.164	0.084	0.041	0.050	0.288	−0.191	0.300	0.024	0.041
	$b = b_{max}$	−0.500	−0.205	−0.152	−0.110	−0.118	−0.330	0.054	0.129	−0.064	−0.066
$n = 256$	$b = 1$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	$b = 2$	0.977	0.595	0.492	0.435	0.372	0.771	0.028	−0.013	0.211	0.183
	$b = 3$	0.953	0.397	0.354	0.265	0.333	0.924	0.391	0.370	0.178	0.145
	$b = 4$	0.931	0.335	0.266	0.224	0.213	0.759	0.015	0.288	0.140	0.140
	$b = b_{max}$	−0.500	−0.156	−0.114	−0.086	−0.096	−0.443	−0.068	−0.032	−0.053	−0.057
$n = 576$	$b = 1$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	$b = 2$	0.990	0.642	0.550	0.498	0.435	0.892	0.098	0.056	0.267	0.233
	$b = 3$	0.979	0.466	0.428	0.344	0.398	0.964	0.425	0.404	0.235	0.197
	$b = 4$	0.969	0.411	0.348	0.307	0.290	0.886	0.085	0.326	0.199	0.191
	$b = b_{max}$	−0.500	−0.137	−0.100	−0.076	−0.086	−0.473	−0.074	−0.049	−0.049	−0.053

$k$  indicates how many immediate neighbours there are. The  $d_{ij}$  are defined by  $d_{ii} = 1/121$ ,  $d_{ij} = (1/k)(120/121)$  for the immediate  $k$  neighbours, and  $d_{ij} = 0$  otherwise.

Moran process and with  $g = 2$  for the Fisher–Wright process. This is seen in the equal small negative values for all  $b$  greater than 1 in the relevant parts of the table. The larger is  $k$ , the lower the tendency to be altruistic to an individual neighbour. Further, we can see the inverted compensated relatednesses with  $g = 1$  in the Fisher–Wright case, where they are most negative to the nearest neighbour, and become less negative for more distant neighbours. In general, the Fisher–Wright relatednesses are lower, with negative values at  $g = 1$ , neutrality at  $g = 2$ , and lower positive values for wider compensation, extending to lower uncompensated relatednesses.

Relatednesses are shown arrayed spatially in Fig. 3. The Fisher–Wright values illustrate the checkerboard effect, with the effects shown graphically in Fig. 4. An individual can share a gene with a neighbour two steps away quite easily, simply by both neighbours being the offspring from the node in between. However, to share a gene with an immediate neighbour, the path of ancestry must include at least one unlikely event—in this particular example, of self-replacement, whose probability is only  $1/121$ . Paths linking immediate neighbours are therefore longer and more indirect, and the result is that an individual is less related to her immediate neighbour than to the neighbour beyond that. In considering the implications of this model for realistic biological situations, it is not impossible that compulsory movers might ‘step round’ each other, but it is most unlikely that such strong separation would arise. Nevertheless, modellers must be aware of these phenomena.

In general, the theory inspired by TDW brings great order to the understanding of these calculations, which are

able to illustrate some important patterns. But it must not be forgotten that the general theory applies to much more general dispersal regimes. There can be different strengths of links, and not just to immediate spatial neighbours, for example. Only a small part of the implications of the general theory can be illustrated with these calculations.

All the calculations were carried out in Mathematica 5.2 (Wolfram Research, Inc., 2005).

#### 4.1. Compensation explained in pictures

The purpose of this section is to clarify that the expression for inclusive fitness using compensated relatednesses is just a convenient shorthand for a much longer expression, based on conventional relatednesses and standard theory, in which each affected individual is identified and included in the formula. Fig. 5 shows an example of how density dependence alters the fitness of individuals surrounding one that suffers an initial fitness change, and identifies them for the sequel.

The nearest neighbours are affected with  $g = 1$  and the neighbours’ neighbours with  $g = 2$ . That includes some twice (because they are the neighbour of more than one of the first circle of neighbours), some just once, and the initial individual herself 4 times, as she is a neighbour of each of her four neighbours. Inclusive fitness theory (e.g. Grafen, 2006) tells us that, writing out each neighbour’s contribution separately, the net inclusive fitness effect for an individual  $i$  of the whole set of changes, assuming the initial individual is  $j = 1$  and the amount of the initial fitness change

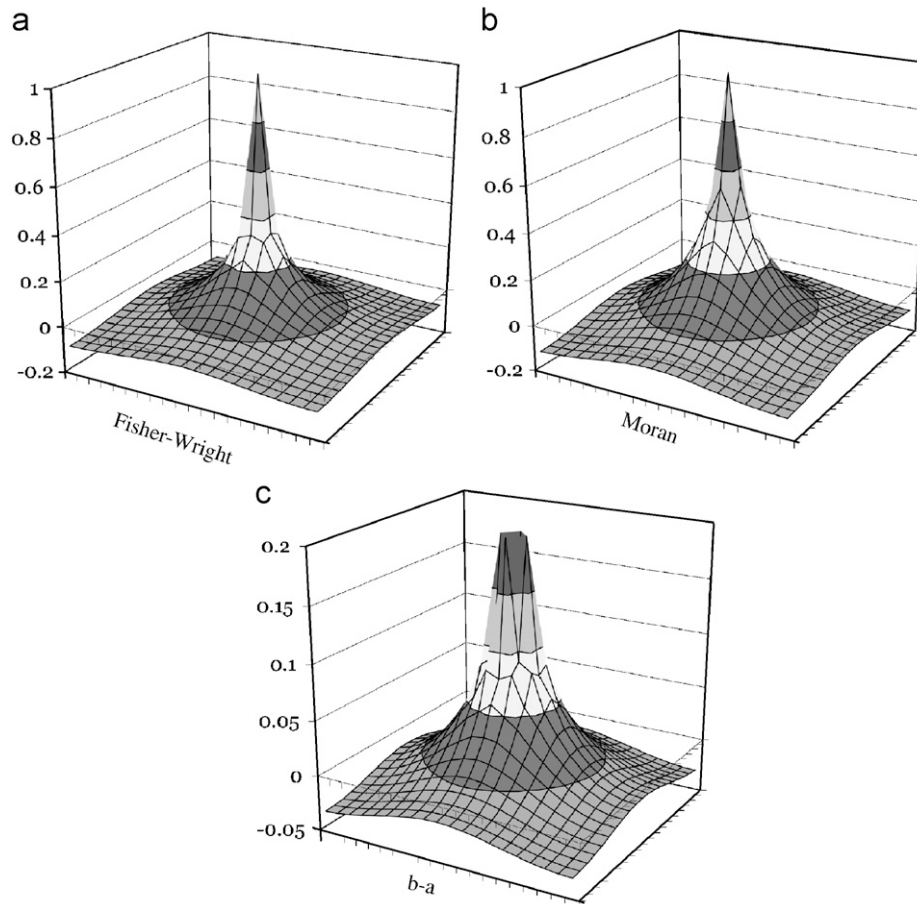


Fig. 2. Relatedness (uncompensated) to the central individual in a square ( $k = 4$ )  $20 \times 20$  grid. It begins at 1, and falls in a witch's hat pattern. The average relatedness is zero, and the low negative relatednesses to distant individuals reflect the higher positive values to a few neighbours. The values are generally more moderate, i.e. closer to zero, under the Fisher–Wright scheme (a) than under the Moran scheme (b). The difference between the Moran and Fisher–Wright relatednesses (c) is zero for the central value, and therefore hidden, but the nearest neighbours form the lip of the crater. All the near neighbours are less related under Fisher–Wright, while the negative differences at the edge show that the negative values are less extreme too than in the Moran case. In calculating the relatedness between two nodes in the Fisher–Wright case, it has to be taken into account that both occupants have been replaced, while in Moran, only one has. The dispersal for this figure assumes  $d_{ij} = 1/5$  for self and each of the four immediate neighbours. This avoids the sawtooth effect of Appendix C, and permits a comparison between the two processes. Note that allowing self-replacement does not affect the relatednesses under the Moran process.

was  $x$ , are

$$r_{i1}x - \left(\frac{1}{4}r_{i2}x + \frac{1}{4}r_{i3}x + \frac{1}{4}r_{i4}x + \frac{1}{4}r_{i5}x\right) \quad (g = 1), \quad (5)$$

$$r_{i1}x - \left(\frac{1}{4}r_{i1}x + \frac{1}{8}r_{i6}x + \frac{1}{16}r_{i7}x + \frac{1}{8}r_{i8}x + \frac{1}{16}r_{i9}x + \frac{1}{8}r_{i10}x + \frac{1}{16}r_{i11}x + \frac{1}{8}r_{i12}x + \frac{1}{16}r_{i13}x\right) \quad (g = 2). \quad (6)$$

As the compensation for  $x$  adds up to a total of  $x$ , we can think of it as suffered by a composite individual, with the average relatedness of the whole circle, where the average must be weighted according to how much of the  $x$  falls on each individual in the circle.

This could be calculated as the ‘economic relatedness’  $r_e$ , which would in these cases be

$$r_e = \left(\frac{1}{4}r_{i2} + \frac{1}{4}r_{i3} + \frac{1}{4}r_{i4} + \frac{1}{4}r_{i5}\right) \quad (g = 1),$$

$$r_e = \left(\frac{1}{4}r_{i1} + \frac{1}{8}r_{i6} + \frac{1}{16}r_{i7} + \frac{1}{8}r_{i8} + \frac{1}{16}r_{i9} + \frac{1}{8}r_{i10} + \frac{1}{16}r_{i11} + \frac{1}{8}r_{i12} + \frac{1}{16}r_{i13}\right) \quad (g = 2).$$

We can then write the inclusive fitness contribution as

$$r_{i1}x - r_e x. \quad (7)$$

Grafen (1984), with some later authors, assumed that the net effect of the action,  $b - c$ , would be compensated together, but Queller (1994) pointed out this is not appropriate if the actor and recipient have different sets of neighbours. The two methods would be equivalent in an island model where altruism was to group-mates, but are different in the more general setting of networks used here, which also includes lattices.

The final step is then natural, to write Eq. (7) as  $(r_{i1} - r_e)x$  and to think of  $r_{i1} - r_e$  as the compensated relatedness for the effect on  $i$ 's inclusive fitness of individual 1 receiving a fitness change of  $x$ . We notate it  $\tilde{r}_{ij(1)}$  or  $\tilde{r}_{ij(2)}$  according as the circle bearing the compensation is 1 or 2. The whole argument works whether  $i$  is 1 or not, though  $r_e$  will be different, and so we can write, for an actor  $i$  losing  $c$  to give

Table 2  
Compensated relatednesses to self in some simple networks, at three different populations sizes, with five different circles of compensation

Population size	Circle	Model									
		Moran					Fisher–Wright				
		<i>k</i> = 2	<i>k</i> = 3	<i>k</i> = 4	<i>k</i> = 6	<i>k</i> = 8	<i>k</i> = 2	<i>k</i> = 3	<i>k</i> = 4	<i>k</i> = 6	<i>k</i> = 8
<i>n</i> = 64	<i>g</i> = 1	0.650	0.633	0.643	0.709	0.787	0.845	1.177	1.220	0.900	0.933
	<i>g</i> = 2	0.331	0.518	0.639	0.703	0.781	0.351	0.430	0.502	0.779	0.840
	<i>g</i> = 3	0.644	0.734	0.789	0.810	0.867	0.840	1.172	1.217	0.894	0.926
	<i>g</i> = 4	0.412	0.679	0.783	0.841	0.894	0.438	0.567	0.619	0.897	0.932
	<i>g</i> = ∞	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
<i>n</i> = 256	<i>g</i> = 1	0.349	0.490	0.504	0.560	0.642	0.428	0.969	1.005	0.782	0.829
	<i>g</i> = 2	0.178	0.406	0.506	0.563	0.645	0.213	0.417	0.502	0.682	0.752
	<i>g</i> = 3	0.347	0.577	0.629	0.653	0.722	0.426	0.969	1.007	0.785	0.832
	<i>g</i> = 4	0.224	0.538	0.630	0.683	0.751	0.267	0.553	0.624	0.792	0.842
	<i>g</i> = ∞	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
<i>n</i> = 576	<i>g</i> = 1	0.238	0.432	0.446	0.498	0.578	0.279	0.899	0.937	0.727	0.778
	<i>g</i> = 2	0.122	0.359	0.449	0.502	0.582	0.142	0.393	0.475	0.634	0.707
	<i>g</i> = 3	0.237	0.510	0.558	0.583	0.653	0.278	0.900	0.940	0.731	0.783
	<i>g</i> = 4	0.153	0.477	0.560	0.611	0.680	0.178	0.521	0.591	0.738	0.793
	<i>g</i> = ∞	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

See Table 1 for further details.

Table 3  
Actor-referenced compensated relatednesses in some simple networks with *n* = 576, with four different circles of compensation *g* = 1, 2, 3, ∞, assuming the same circle for actor and recipient

Circle	Neighbour	Model									
		Moran					Fisher–Wright				
		<i>k</i> = 2	<i>k</i> = 3	<i>k</i> = 4	<i>k</i> = 6	<i>k</i> = 8	<i>k</i> = 2	<i>k</i> = 3	<i>k</i> = 4	<i>k</i> = 6	<i>k</i> = 8
<i>g</i> = 1	<i>b</i> = 1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	<i>b</i> = 2	−0.002	−0.002	−0.002	−0.002	−0.002	−0.834	−0.574	−0.506	−0.136	−0.084
	<i>b</i> = 3	−0.002	−0.002	−0.002	−0.002	−0.002	0.694	0.375	0.376	0.043	−0.119
	<i>b</i> = 4	−0.002	−0.002	−0.002	−0.002	−0.002	−0.579	−0.317	0.300	0.010	0.037
	<i>b</i> = <i>b</i> <sub>max</sub>	−0.002	−0.002	−0.002	−0.002	−0.002	−0.000	0.003	0.009	−0.001	−0.001
<i>g</i> = 2	<i>b</i> = 1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	<i>b</i> = 2	0.490	0.326	0.243	0.161	0.120	−0.002	−0.002	−0.002	−0.002	−0.002
	<i>b</i> = 3	−0.003	−0.003	−0.003	−0.003	0.120	−0.002	−0.002	−0.002	−0.002	−0.002
	<i>b</i> = 4	−0.003	−0.003	−0.003	−0.003	−0.003	−0.002	−0.002	−0.002	−0.002	−0.002
	<i>b</i> = <i>b</i> <sub>max</sub>	−0.003	−0.003	−0.003	−0.003	−0.003	−0.002	−0.002	−0.002	−0.002	−0.002
<i>g</i> = 3	<i>b</i> = 1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	<i>b</i> = 2	0.334	0.249	0.198	0.186	0.162	−0.743	−0.428	−0.379	0.007	0.027
	<i>b</i> = 3	0.161	0.078	0.094	0.042	0.135	0.693	0.373	0.374	0.041	−0.008
	<i>b</i> = 4	−0.003	−0.004	0.045	0.019	0.036	−0.579	−0.317	0.298	0.008	0.035
	<i>b</i> = <i>b</i> <sub>max</sub>	−0.003	−0.004	−0.004	−0.004	−0.005	−0.000	0.003	0.008	−0.002	−0.002
<i>g</i> = ∞	<i>b</i> = 1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	<i>b</i> = 2	0.573	0.381	0.305	0.233	0.199	0.003	0.002	0.001	0.046	0.054
	<i>b</i> = 3	0.163	0.079	0.095	0.062	0.172	0.163	0.080	0.096	0.044	0.026
	<i>b</i> = 4	0.076	0.049	0.045	0.039	0.054	−0.002	−0.003	0.047	0.021	0.038
	<i>b</i> = <i>b</i> <sub>max</sub>	−0.005	−0.005	−0.006	−0.006	−0.006	−0.002	−0.003	−0.003	−0.003	−0.003

For further details see Table 1.

*b* to individual *j*, the inclusive fitness effect of the act as

$$\tilde{r}_{ij(g)}b - \tilde{r}_{ii(g)}c$$

thus obtaining the desired endpoint.

We end this section by explaining how easily the critical cost–benefit ratio can be calculated numerically. To understand under which conditions altruism can evolve we need to know the dispersal probabilities (*d*<sub>*ij*</sub>), which are

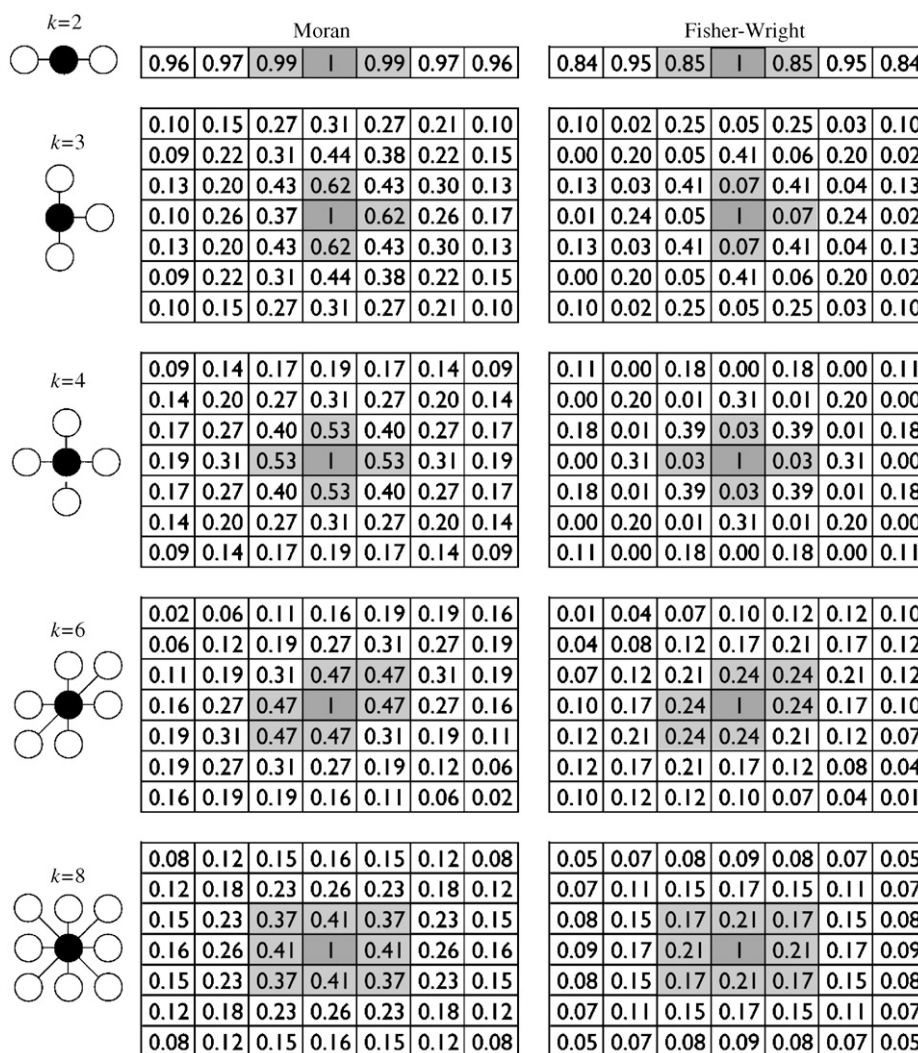


Fig. 3. The figure shows the relatednesses to the central individual as arranged on the grid. Only the central part is shown, so the negative relatednesses to distant individuals apparent in Fig. 2 do not appear here. The grid was 20 × 20, and it was assumed that each individual had four immediate neighbours. The calculations assumed a 1/121 chance of self-replacement.

in principle estimable from field data, and the compensation scheme ( $g$ ), which, as explained before, will depend on the kind of behaviour we are trying to understand. Suppose we want to calculate the cost–benefit ratio that would be favoured by selection, under the Moran scheme when compensation for both actor and recipient is in the 2-circle, in a square grid where each individual has four immediate neighbours, with a population size of 576. Let us suppose the dispersal scheme of the tables which has  $d_{ii} = 1/121$  and  $d_{ij} = 30/121$  for the neighbours. The actor-referenced compensated relatedness can be calculated from Eq. (4) by dividing relatedness to the neighbour  $(30/121 - (2 - 30/121)/575)$  by relatedness to self  $(1 + 1/121 - (1 - 1/121)/575)$ , obtaining  $8519/35015 = 0.243296$ . This confirms the value of 0.243 from Table 3, and we conclude from Eq. (1) that inclusive fitness has the same sign as  $0.243296b - c$ . Altruism to an immediate neighbour will therefore evolve if  $b/c > 4.11022$ .

#### 4.2. The differences between DB and BD explained

It is a curious but reliable feature of results of a number of authors (Ohtsuki et al., 2006; Ohtsuki and Nowak, 2006; Grafen, 2007; Lehmann et al., 2007a) that neighbourly altruism evolves very differently under so-called DB and BD demographic schemes, using the Moran replacement scheme in which one individual dies and is replaced at each time step. TDW explain, though not in these terms, that the circle of compensation is the 1-circle in BD schemes, but the 2-circle in DB schemes. As the compensation is in the balancing circle under BD, no altruism evolves, but as compensation is spread more widely there, altruism does evolve under the DB scheme.

The question considered in this section is: *why* does the demographic scheme affect the circle of compensation? The difference between the schemes is apparently whether birth or death comes first, but can this be biologically significant?



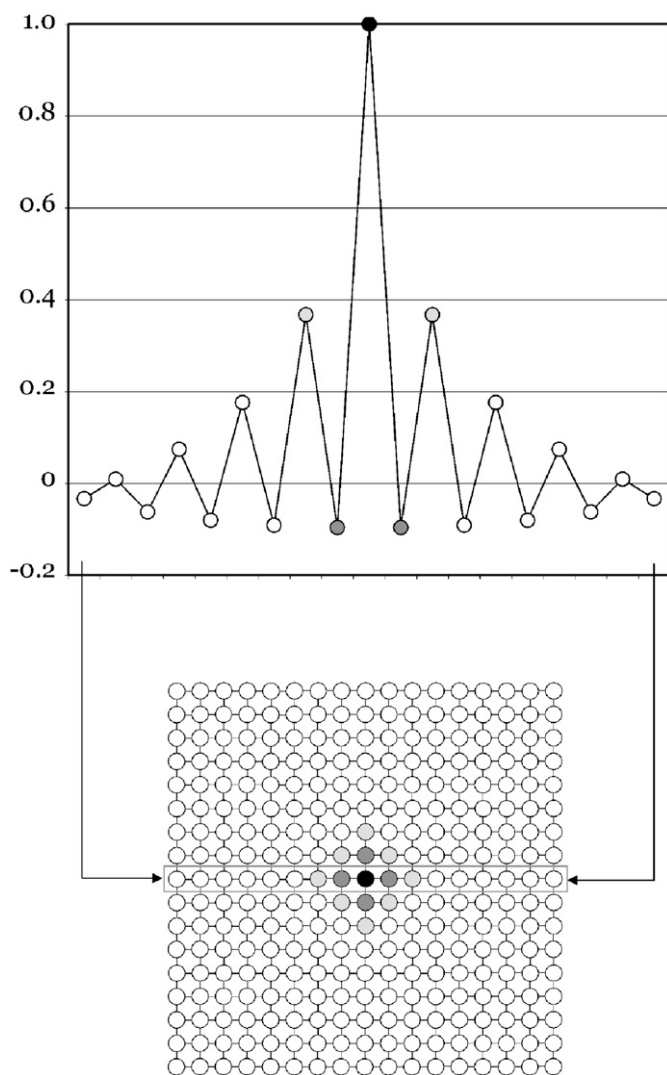


Fig. 4. The figure shows how relatednesses oscillate from the central individual to succeeding neighbours. This can occur in Fisher–Wright processes—see Appendix C.

There are two distinctions to be drawn about a stage in the demographic scheme. A stage can be global or local, according as the decisions made in that stage select one from the population, or one from a small group. It is convenient to use the word ‘score’ to refer to the fitness effects prior to the operation of density dependence, so that the altruistic act subtracts  $c$  from the actor’s score and adds  $b$  to the recipient’s score. A stage in the demographic scheme can be random or it can be ‘score dependent’. All the authors listed made the first stage global, and the second stage local, and further they all made B score dependent and D global. To investigate this phenomenon, we constructed models in which death but not birth was score dependent, and in which both birth and death were score dependent. In order to maintain the sense that a higher score is an advantage, and relying on the weak selection assumption to prevent scores ever becoming zero or negative, we assumed that where the chance of death did depend on score, it was *inversely* proportional to the score.

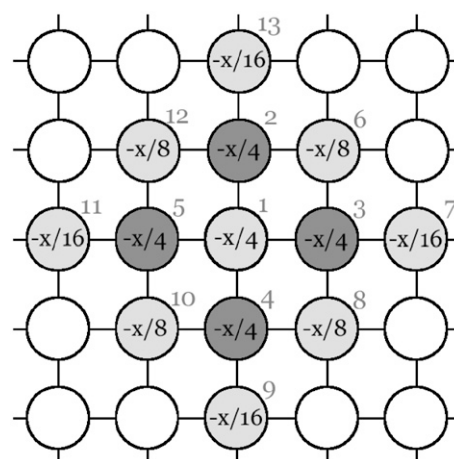


Fig. 5. The figure shows how the density-dependent compensation is spread around the individual with the initial fitness effect of size  $x$ . Inside each dark-shaded circle is shown the compensation experienced when  $g = 1$ , to the immediate neighbours. In each light-shaded circle is shown the compensation experienced when  $g = 2$ , and note that the individual herself suffers a quarter of it. The numbers just outside the circles identify the individuals for the purposes of Eqs. (5) and (6).

Calculations of the weak selection fitness effects showed that the D(random) – B(score) scheme previously assumed is identical to our new B(random) – D(score) scheme, and that the previous B(score) – D(random) scheme is identical to our new D(score) – B(random). Both BD and DB schemes in which both stages were score dependent allowed altruism to evolve under the exact same conditions as each other, but with roughly double the critical  $c/b$  ratio required where only one stage was score dependent. Thus the key issue is whether score affects only local competition, in which case local altruism can evolve; or only global competition, in which case it does not; or both, in which altruism can evolve but with more difficulty.

It is clear that it is only the score-dependent stage or stages that mediate the effect of altruism. Score dependence in global competition is compensated in the 1-circle, and so is virtually neutral ( $\tilde{r}$  to neighbours is  $-1/(n-1)$ , Moran case) or positively disadvantageous ( $\tilde{r}$  is inverted, and negative, Fisher–Wright). Score dependence in local competition cases is compensated in the 2-circle, and so altruism can evolve ( $\tilde{r}$  is positive, Moran) or not ( $\tilde{r}$  to neighbours is  $-1/(n-1)$ , Fisher–Wright).

Thus the interaction of local and global competition versus score dependence and non-score dependence may be a biologically meaningful explanation for the difference between the demographic schemes. The order of birth and death, however, is not.

### 4.3. Redefinitions of altruism

The consistent use of the term altruism has been powerfully advocated by Rousset (2004, p. 106), as he endorses ‘the dour conclusion... that much of the debate about ‘group selection’ and ‘altruism’ boils down to an

inconsistent use of words'. In this paper, it has been found convenient to use altruism in relation to the pre-density-dependent effects, contrary to Rousset's advice. The diagram of the previous section is convenient for discussing this possible redefinition of altruism. The original 'Hamilton definition' of altruism is an action with a positive cost to the actor and a positive benefit to the recipient, both measured as numbers of offspring. This must mean the net number of offspring after taking density-dependent compensation into account, as endorsed by Rousset (2004).

It is necessary to set up a notation in which the redefinition can be explored. We have used  $b$  and  $c$  for the pre-density-dependence benefit and cost of an action. Let  $\tilde{b}$  and  $\tilde{c}$  represent the post-density-dependence benefit and cost of the action to the recipient and actor, and let  $\tilde{B}$  represent the summed post-density-dependence benefit to all other individuals in the population. Node transitivity implies that each individual loses the same fraction of its original fitness effect through density dependence, and let this fraction be  $\mu$ ; and together with symmetry of the dispersal matrix, it implies that the actor and recipient are affected by the same fraction of each other's original fitness effect, which we shall call  $\lambda$ .  $\mu = 1/4$  in the example, but  $\lambda$  is not specified, as it depends on how close the actor and recipient are. If the actor were individual number 10, then  $\lambda = 1/8$ .

These definitions allow us to write

$$\tilde{c} = c - \mu c + \lambda b,$$

$$\tilde{b} = b - \mu b + \lambda c,$$

$$\tilde{B} = (b - c)(1 - \mu - \lambda)$$

and we note that as required by inelastic density dependence,  $-\tilde{c} + \tilde{b} + \tilde{B} = 0$ .

If an action is altruistic according to the original  $b$  and  $c$ , then  $b, c > 0$ , and this is easily seen to imply that  $\tilde{b}, \tilde{c} > 0$ . Thus any action that is altruistic before density dependence remains altruistic afterwards. The main problem in the literature arises from the dubious distinction between weak and strong altruism, in which an action is selected that is altruistic according to the author, but not according to the original Hamilton definition. Here, this cannot happen because all altruistic acts taking the pre-density-dependent values for  $b$  and  $c$  are also altruistic in Hamilton's sense. The distinction from the usual case of concern is that there, the actor's help is directed to a group that includes itself, and so the benefit to self is greater than originally appears. Here, the actor's help is directed at a recipient, and the density-dependent effects on an actor of a positive benefit to the recipient are negative.

With this methodological worry assuaged, we can enquire whether it is helpful to use the term altruism in relation to the original values of  $b$  and  $c$ . It is usual in theory to sweep density dependence under the carpet, and assume it happens 'out of sight'. It is probably true that calculating empirically the effect of an action, taking out the density-dependent effects, has never been done, and

would be rather difficult, involving observations for much longer periods and extending over many more phenomena, than simply measuring an initial fitness change in some currency.

To insist on using the Hamilton sense would also be theoretically complicated, because there are many more recipients than one—all individuals affected by the density-dependent compensation are also recipients of the action on that analysis. Altruism is defined only dyadically, or at least with a number of equal recipients, and we would require a definition that covered a situation in which there was an actor suffering a cost, a primary recipient gaining a benefit, and several secondary recipients, suffering varying positive and negative fitness effects.

Our tentative conclusion is that it is worthwhile, and actually methodologically extremely convenient, to apply the definitions of altruism to the pre-density-dependent effects, although stressing it is important always to bear in mind in drawing conclusions that this definition is not the original Hamilton definition.

A final point concerns the difference between altruism and cooperation. Altruism seems to us the preferable term for the trait we are studying, because it is unilaterally decided. Others, including TDW, have used the term 'cooperation', but we feel this should be reserved for biological situations in which the effect of one individual's decision depends on the actions of others. Pointedly, consider the case in which an allele is expressed with a penetrance of only 0.1. The theory says the direction of selection is unchanged, which is quite appropriate when there is no synergy in the actions of the individuals. But in a paired case, a penetrance of 0.1 would leave only 0.01 of pairs showing the 'cooperation' in both individuals, and 0.18 of pairs having 'cooperation' in only one individual. If the selective effect of an individual's action does not depend on the behaviour of the other, this does not deserve the 'co' in 'cooperation'.

## 5. Discussion

The theory of TDW, and as extended here, is appealingly general within prescribed mathematical conditions, covering all graphs subject to node transitivity and symmetry of the dispersal matrix. The population is assumed to be finite, which is not limiting biologically, but does restrict the models to which the theory can be applied. Subject to finiteness of population, the range of models covered includes all the Wrightian island and stepping stone models, and hierarchical versions of them, and lattice models following Malécot, as well as the regular graphs of evolutionary graph theory. TDW present a 'zoo' of such graphs, and to them we can add the graphs included by the weakening of assumptions in Appendix B.

There are various biological limitations. The theory assumes haploidy and asexuality, whereas the diploid sexual case would be more relevant to most studied organisms, and technical challenges will be posed by the

usual sex differences in dispersal pattern (Greenwood, 1980). A further limitation is homogeneity of the population, embodied in the assumption of node transitivity, whose most significant implication is that there are no sinks or sources in the population, and also that none of the nodes is permitted to be unoccupied. Symmetry of the dispersal matrix is also required, which might be untrue if prevailing winds or currents made dispersal easier in one direction than another.

A major limitation is inelasticity of the environment and the population. Changing traits of the population cannot increase or decrease the population size, or local group size, and this limits the kind of social traits that can plausibly be represented. Elastic environments (Lehmann et al., 2006) are a more generally realistic case.

The life-cycle assumptions are general in that they include the Fisher–Wright process at one end, and limit to the Moran process (at least for present purposes) at the other. However, there limitations too. In the Moran process, one individual is replaced in each time step. The analytical results assume that the altruistic behaviour affects *either* the chance of being the replaced individual *or* the chance of providing the offspring to replace that individual, though we considered the combination in the numerical calculations of Section 4. Further, when more than one individual is replaced in a time-step, we have implicitly made the assumption that, conditional on which individuals are to be replaced, the processes of replacement are independent. This amounts to an assumption of very many offspring produced in each time-step, so that if I supply the replacement for one individual that does not affect my chances in another replacement.

The value of the theory developed here, extending the important results of TDW, is that it provides a very powerful and explicit analysis of a well-defined general case, which may prove to be useful in further developing our understanding of more realistic cases.

### Acknowledgements

The paper is a development of important insights of Taylor et al. (2007a), and we are very grateful to Peter Taylor, Troy Day and Geoff Wild for sight of the manuscript of that paper. Laurent Lehmann, Laurent Keller and David Sumpter generously allowed us sight of a MS of Lehmann et al. (2007a). We are grateful to the authors of both those papers, and to Martin Nowak, for helpful discussions about inclusive fitness on graphs. Laurent Lehmann, Peter Taylor, Andy Gardner, Stu West and Laurent Keller made extremely useful comments on the manuscript. The counter-example in Fig. B1 was pointed out to us by Colin McDiarmid and Peter Cameron, who kindly gave us permission to include it here, and we are grateful to Philip Maini for passing the problem on to them. The very neat formulation for compensated relatednesses in Eq. (4) and following equations was acutely and generously suggested by one referee. The other very rightly

encouraged a better attempt to make connections to the literature on structured populations that predates the graph theory approach. We are grateful to them both. M.A. is supported by a long-term fellowship of the Human Frontier Science Program Organization.

### Appendix A. Theoretical development

Each of a finite set of  $n$  nodes is occupied by one individual, and the individuals reproduce and die through a series of steps. The probability that the parent of an individual at node  $i$  occupied node  $j$  is  $D_{ij}$ . The matrix representation in this appendix will use capital letters for variables represented by lower case letters in the main text. Deaths occur at a series of steps, and at each step, each individual has a (pairwise independent) probability  $s$  of surviving, and  $1 - s$  of dying and being replaced. Each offspring has an independent probability  $\mu$  of mutating to a new genotype. Let  $(G_t)_{ij}$  be the probability that individuals at nodes  $i$  and  $j$  have genotypes that are identical by descent at step  $t$ . Then we obtain the following equation for  $G_{t+1}$ :

$$G_{t+1} = s^2 G_t + 2s(1-s)\left(\frac{1}{2}(1-\mu)(DG_t + G_t D^T)\right) + (1-s)^2(1-\mu)^2 DG_t D^T + L_{t+1},$$

where  $L_{t+1}$  is a diagonal matrix that ensures the diagonal elements of  $G_{t+1}$  each equal 1.

It may be helpful to remind readers of the definitions of Gantmacher (1960), as we will use his terminology and rely on theorems mainly in Volume 2, Chapter XIII, Section 7. A stochastic matrix is regular if all its eigenvalues of modulus 1 are equal to 1, and fully regular if there is only one such eigenvalue and it is algebraically simple.

We now assume that  $D$  is symmetric, and without loss of generality we declare all alleles to be non-identical by descent in generation 0, so  $G_0 = I$ . We also assume node transitivity, which guarantees that the diagonal elements of  $L_t$  are all equal to each other. On this basis it easy to show that  $G_{t+1}$  commutes with  $D$  provided  $G_t$  does. As  $G_0 = I$ , we conclude that  $G_t$  commutes with  $D$  for all  $t$ . Hence we can define bistochastic matrix  $E$  and scalar  $e > 0$  by

$$eE = s^2 I + 2s(1-s)(1-\mu)D + (1-s)^2(1-\mu)^2 D^2, \\ E\mathbf{1} = \mathbf{1}$$

and then define  $\eta_t$  by

$$G_{t+1} = eEG_t + \eta_{t+1}I, \\ (G_t)_{ii} = 1$$

and so ‘solve’ the recursion as

$$G_t = \sum_{m=0}^{t-1} \eta_{t-m} e^m E^m.$$

We assume that  $G_t$  converges to  $G_\infty(\mu)$ , say, and so that  $\eta_t$  does too, to  $\eta_\infty(\mu)$ , and that both limits are continuous functions of  $\mu$ . If  $E$  is regular, then  $E^\infty$  exists and is proportional to the limit of  $G_\infty$  as  $\mu \rightarrow 0$ . If  $E$  is fully

regular,  $E^\infty = \mathbf{1}\mathbf{1}^\top/n$  and so  $G_\infty = \mathbf{1}\mathbf{1}^\top$ . We assume full regularity of  $E$  from now on, and so exclude the possibilities that the network is not connected, and that there are separate non-intersecting subpopulations.

It is worth noting that these assumptions exclude some examples studied in the literature. For example, consider the square grid when  $d_{ij} = 1/4$  for four immediate neighbours from Fig. 1. When  $s = 0$ , we have  $E = D^2$ , and  $D^2$  is reducible, reflecting the complete separation of two subpopulations corresponding to the white and black squares on a chessboard. Such difficulties have been avoided in the numerical examples of Section 4 by allowing a small chance of self-replacement.

We proceed to find the relatednesses in the limit as  $\mu \rightarrow 0$ , by dropping the time-step subscripts on  $G$ , differentiating with respect to  $\mu$ , evaluating at  $\mu = 0$ , dividing by  $1 - s$ , and solving for  $\eta' = 2n(1 - s)$  (by post-multiplying by  $\mathbf{1}$ ) to obtain

$$(1 + s)G' = 2s(1 - s)DG' + (1 - s)^2D^2G' - 2(1 - s)\mathbf{1}\mathbf{1}^\top + 2nI.$$

We now show that the relatedness matrix  $R$  is equal to  $\mathbf{1}\mathbf{1}^\top + \alpha^{-1}G'$  where  $\alpha$  is defined by  $R\mathbf{1} = 0$ . Relatedness is defined, for example by TDW, in terms of identity by descent as  $r_{ij} = (g_{ij} - \bar{g})/(1 - \bar{g})$  where  $\bar{g}$  is the average of the  $g_{ij}$ . All the  $g_{ij}$  go to 1 as  $\mu \rightarrow 0$ , so we rewrite and use L'Hôpital's Rule as follows:

$$1 - r_{ij} = \frac{\lim_{\mu \rightarrow 0} (1 - g_{ij})/\mu}{\lim_{\mu \rightarrow 0} (1 - \bar{g})/\mu},$$

$$1 - r_{ij} = \alpha^{-1} \lim_{\mu \rightarrow 0} \frac{1 - g_{ij}}{\mu},$$

where  $\alpha$  is defined as the denominator. The remaining limit equals  $-g'_{ij}$ , thus showing

$$r_{ij} = 1 + \alpha^{-1}g'_{ij}$$

which translated into matrix notation gives the required result. The requirement  $R\mathbf{1} = 0$  ensures that the mean relatedness to the population is zero.

Proceed by substituting for  $G'$  to obtain

$$(1 + s)R = 2sDR + (1 - s)D^2R - 2\alpha\mathbf{1}\mathbf{1}^\top + 2\alpha nI$$

and, substituting  $\lambda = 2\alpha/(1 + s)$ ,

$$R = \left( \frac{2s}{1+s}D + \frac{1-s}{1+s}D^2 \right) R - \lambda\mathbf{1}\mathbf{1}^\top + \lambda nI, \quad (8)$$

where we can conveniently define  $\lambda$  as maintaining the diagonal elements of  $R$  at 1.

Defining the factor of  $R$  in Eq. (8) as  $C$ , we can write

$$R = CR - \lambda\mathbf{1}\mathbf{1}^\top + \lambda nI.$$

$C$  is an average of  $D$  and  $D^2$ , and so also commutes with  $R$ . The relatedness of individual  $i$  to the 1-circle (resp.  $g$ -circle) of individual  $j$  is given by the  $i, j$ th element of  $DR$  (resp.  $D^gR$ ). The relatedness of individual  $i$  to the weighted mixture  $2s/(1 + s)$  of the 1-circle and  $(1 - s)/(1 + s)$  of the

2-circle is given by  $CR$ , and to the  $k$ th self-convolution of that mixture by  $C^kR$ .

The compensated relatedness to the  $k$ th self-convolution is  $\tilde{R}_{(\tau^k)} = R - C^kR$ . If  $C$  is fully regular, the limit  $C^\infty = \mathbf{1}\mathbf{1}^\top/n$ , and we now make that assumption for the remainder of the appendix. It is easy to confirm that

$$\tilde{R}_{(\tau^k)} = \lambda \sum_{m=0}^{k-1} (nC^m - \mathbf{1}\mathbf{1}^\top). \quad (9)$$

(If  $C$  is not fully regular, then compensation is not spread evenly as  $k$  increases, but asymptotically is concentrated on particular subsets.) The simplest case of  $k = 1$  shows that

$$\lambda^{-1}\tilde{R}_{(\tau)} = nI - \mathbf{1}\mathbf{1}^\top \quad (10)$$

which shows that the relatednesses are in the same proportions (1 to self and  $-1/(n - 1)$  to each other individual) as in the non-viscous case. We define  $r_\tau$ , the relatedness to the  $\tau$ -circle, as the common diagonal element of  $CR$ , and note that it satisfies

$$r_\tau = 1 - \lambda(n - 1), \quad \lambda = \frac{1 - r_\tau}{n - 1}.$$

At the other extreme, the limit of  $R - C^kR$  as  $k \rightarrow \infty$  exists and is equal to  $R$  itself, because  $C^\infty R = \mathbf{1}\mathbf{1}^\top R/n$  and  $\mathbf{1}^\top R = 0$ . This provides a formula for  $R$

$$R = \tilde{R}_{(\tau^{\infty})} = \lambda \sum_{m=0}^{\infty} (nC^m - \mathbf{1}\mathbf{1}^\top)$$

and it is worth recording the inverse relationship

$$C = I - \lambda \sum_{m=0}^{\infty} (n(I - R)^m - \mathbf{1}\mathbf{1}^\top) \quad (11)$$

which shows how to derive  $C$  from  $R$ . The value of  $\lambda$  can be expressed in terms of the eigenvalues  $\gamma_n$  of  $C$  as

$$\lambda = \frac{1}{\sum_{n \neq t} 1/(1 - \gamma_n)}, \quad (12)$$

where  $t$  indexes the leading eigenvalue  $\gamma_t = 1$  corresponding to the eigenvector  $\mathbf{1}$ . On the assumption of full regularity there is exactly one eigenvalue of modulus 1 and it equals 1.

## Appendix B. Assumptions on $D$

A permutation of the nodes can be represented as an invertible matrix  $P$ ,  $P\mathbf{1} = \mathbf{1}$  and  $P_{ij} \in \{0, 1\}$ , so also  $P^{-1} = P^\top$ . Let  $u_i$  be the vector with a 1 in position  $i$  and zeroes elsewhere. Then we can express the assumption of node bitransitivity, introduced in this context by Taylor et al. (2007a), as

For each ordered pair of nodes  $(i, j)$  there exists a permutation  $P$  such that  $Pu_i = u_j$ ,  $Pu_j = u_i$  and  $PDP^\top = D$ .



The significant implications are that  $D$  and  $DG$  are symmetric, from which it follows that  $D$  and  $G$  (and so  $G'$  and  $R$ ) commute.

The conclusion that  $G$  and  $D$  commute is reached in Appendix A on the basis of the weaker assumptions of symmetry of  $D$ , and of node transitivity—which states that for each ordered pair of nodes  $(i, j)$  there is a permutation such that  $Pu_i = u_j$  and  $PDP^T = D$ . Indeed the only property for which node transitivity is required is that the diagonal elements of each power of  $D$  are equal to each other (though not necessarily equal to those for different powers), which we term ‘the diagonals condition’.

We record two graphs that are included in the weaker assumptions but are excluded by node bitransitivity. One is graph (e) in Box 1 in TDW itself. They point out that although not satisfying node bitransitivity, it does conform to their conclusions, and our weakening of the assumptions explains why. The other is a cycle with an even number of nodes, in which the  $D_{ij}$  between neighbours alternates around the cycle.

Colin McDiarmid and Peter Cameron kindly pointed out to us a well-known kind of graph that fulfils the diagonals condition but does not satisfy node transitivity. A particular example has 25 nodes, and each node is connected to 12 others. Arranged in a five-by-five grid, with a Latin square arrangement as shown in Fig. B1, each node is connected to all nodes in the same row, and in the same column, and that share the same letter.

It is easy to see that node transitivity is lacking. The largest complete subgraphs are of order 5, and each node belongs to three of them: one each for the row, column and letter. There are also complete subgraphs of order 4 that are not subsets of those of order 5, and they are seen in the

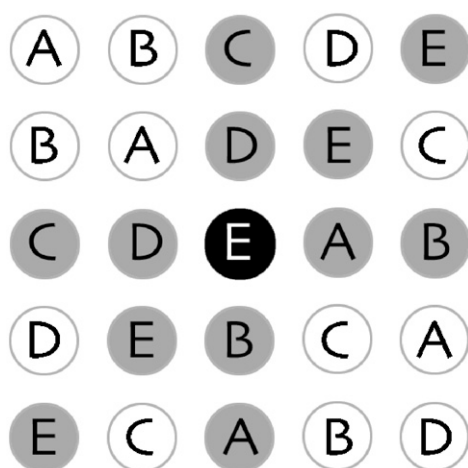


Fig. B1. The figure describes a graph with 25 nodes that is not node-transitive, but does have the property that for each of the powers of the dispersal matrix, the diagonal elements are all equal. A node is connected to all nodes (i) in the same row (ii) in the same column or (iii) sharing the same letter. This is illustrated for the central ‘E’ in black, with all its connected nodes shown in grey. Each node has 12 neighbours, and each of those 12 has an equal chance of providing the replacement offspring when the node’s occupant dies.

figure as four nodes that are at the vertices of a rectangle that has the same letter at opposite vertices. The top-left 2 by 2 corner is an example. The number of such complete 4-subgraphs to which the elements belong varies from 4 (the B in the first column), to one (all other elements in the first column or second row, and four further elements) to zero (12 nodes, including the D at bottom right). Node transitivity is thus impossible, as the number of complete 4-subgraphs to which an element belongs has to be invariant under isomorphisms.

The diagonals condition, however, is met. Let the adjacency matrix be  $A$ . The graph is *strongly regular* and it is a textbook result that therefore  $A^2$  is a linear combination of  $I$ ,  $A$  and  $\mathbf{1}\mathbf{1}^T$  (e.g. Bollobás, 1998, p. 274). Because  $A\mathbf{1}/12 = \mathbf{1}$ , it follows by induction that  $A^k$  is also a linear combination of  $I$ ,  $A$  and  $\mathbf{1}\mathbf{1}^T$ , and from this it follows that the diagonal elements of  $A^k$  are all equal. The dispersal matrix  $D$  is  $A/12$ , and so this graph meets the diagonals condition.

We do not want to replace the condition on node transitivity with the condition of strong regularity, however, as node transitivity permits weighted links between nodes, whereas strong regularity requires two nodes to be either connected or not. The best we can say is that node transitivity and symmetry are jointly sufficient for commutativity of  $D$  and  $G$ , but not necessary. It remains to be seen whether the extra freedom implied by the existence of this counter-example can be turned to biological use.

### Appendix C. The sawtooth effect

In the Fisher–Wright ( $s = 0$ ) case, balancing compensation occurs at the 2-circle. It is to be expected on general grounds that when  $g = 1$ , compensated relatednesses are negative for neighbours, and more negative for closer neighbours. This phenomenon is demonstrated in the numerical examples in Section 4: here it is shown that there are implications for the pattern of compensated relatednesses under the Fisher–Wright process. Note that Appendix B discusses compensation in the  $\tau$ -circle and in circles of the  $k$ th self-convolution of  $\tau$ . Here we discuss compensation in the  $g$ -circles for integer  $g$ . In the terms of the Fisher–Wright case, the previous appendix discussed  $\tilde{R}_{(\tau^{*k})} = R - C^k R$ , while here we discuss  $\tilde{R}_{(g)} = R - D^g R$ . It follows from this definition that

$$\tilde{R}_{(g+1)} = D\tilde{R}_{(g)} + \tilde{R}_{(1)},$$

$$\tilde{R}_{(g)} = \left( \sum_{m=0}^{g-1} (D - \mathbf{1}\mathbf{1}^T/n)^m \right) \tilde{R}_{(1)}.$$

The subtraction of  $\mathbf{1}\mathbf{1}^T/n$  does not affect the product with  $R$  as  $\mathbf{1}^T R = 0$  and  $\mathbf{1}^T D = \mathbf{1}^T$ , and it allows the matrix powers to converge to the zero matrix. The sawtooth emerges when  $D$  has an eigenvalue close to  $-1$ , as the odd and even powers of the matrix pull the relatednesses in opposite directions. Numerical examples are shown in Section 4, and the sawtooth is illustrated in Fig. 4.

## References

- Bollobás, B., 1998. *Modern Graph Theory*. Springer, New York.
- Ewens, W.J., 2004. *Mathematical Population Genetics I. Theoretical Introduction*. Springer, Berlin, Heidelberg, New York.
- Fontanillas, P., Petit, E., Perrin, N., 2004. Estimating sex-specific dispersal rates with autosomal markers in hierarchically structured populations. *Evolution* 58, 886–894.
- Gantmacher, F.R., 1960. *The Theory of Matrices*, vol. 2. Chelsea Publishing Company, New York, NY.
- Gardner, A., West, S.A., 2004. Spite and the scale of competition. *J. Evol. Biol.* 17, 1195–1203.
- Gardner, A., West, S.A., 2006. Demography, altruism, and the benefits of budding. *J. Evol. Biol.* 19, 1707–1716.
- Gardner, A., West, S.A., Buckling, A., 2004. Bacteriocins, spite and virulence. *Proc. R. Soc. Ser. B* 271, 1529–1535.
- Grafen, A., 1984. Natural selection, kin selection and group selection. In: Krebs, J.R., Davies, N.B. (Eds.), *Behavioural Ecology*, second ed. Blackwell Scientific Publications, Oxford, UK, pp. 62–84.
- Grafen, A., 1985. A geometric view of relatedness. *Oxford Surv. Evol. Biol.* 2, 28–89.
- Grafen, A., 2006. Optimisation of inclusive fitness. *J. Theor. Biol.* 238, 541–563.
- Grafen, A., 2007. Inclusive fitness on a cyclical network. *J. Evol. Biol.* 20, 2278–2283.
- Greenwood, P., 1980. Mating systems, philopatry and dispersal in birds and mammals. *Anim. Behav.* 28, 1140–1162.
- Hamilton, W.D., 1964. The genetical evolution of social behaviour. *J. Theor. Biol.* 7, 1–52.
- Hamilton, W.D., 1970. Selfish and spiteful behaviour in an evolutionary model. *Nature* 228, 1218–1220.
- Hamilton, W.D., 1971. Selection of selfish and altruistic behaviour in some extreme models. In: Eisenberg, J., Dillon, W. (Eds.), *Man and Beast: Comparative Social Behavior*. Smithsonian Press, Washington, DC, pp. 57–91.
- Irwin, A., Taylor, P.D., 2001. Evolution of altruism in stepping-stone populations with overlapping generations. *Theor. Popul. Biol.* 60, 315–325.
- Kelly, J.K., 1994. The effect of scale dependent processes on kin selection: mating and density regulation. *Theor. Popul. Biol.* 46, 32–57.
- Lehmann, L., Balloux, F., 2007. Natural selection on fecundity variance in subdivided populations: kin selection meets bet hedging. *Genetics* 176, 1–17.
- Lehmann, L., Perrin, N., Rousset, F., 2006. Population demography and the evolution of helping behaviors. *Evolution* 60, 1137–1151.
- Lehmann, L., Keller, L., Sumpter, D., 2007a. The evolution of helping and harming on graphs: the return of the inclusive fitness effect. *J. Evol. Biol.* 20, 2284–2295.
- Lehmann, L., Roze, D., Rousset, F., Keller, L., 2007b. Strong reciprocity or strong ferocity? A population genetic view of the evolution of altruistic punishment. *Am. Nat.* 170, 21–36.
- Malécot, G., 1948. *Les Mathématiques de l'hérédité*. Masson, Paris (Translated by Demetrios M. Yermanos into English as 'The Mathematics of Heredity', 1969, W.F. Freeman, San Francisco).
- Malécot, G., 1950. Quelques schémas probabilistes sur la variabilité des populations naturelles. *Ann. Univ. Lyon A* 13, 37–60.
- Ohtsuki, H., Nowak, M., 2006. Evolutionary games on cycles. *Proc. R. Soc. Ser. B* 273, 2249–2256.
- Ohtsuki, H., Hauert, C., Lieberman, E., Nowak, M., 2006. A simple rule for the evolution of cooperation on graphs and social networks. *Nature* 441, 502–505.
- Queller, D.C., 1994. Genetic relatedness in viscous populations. *Evol. Ecol.* 8, 70–73.
- Queller, D.C., Goodknight, K., 1989. Estimation of relatedness from allozyme data. *Evolution* 43, 258–275.
- Rousset, F., 2004. *Genetic Structure and Selection in Subdivided Populations*. Princeton University Press, Princeton, NJ.
- Rousset, F., Billiard, S., 2000. A theoretical basis for measures of kin selection in subdivided populations: finite populations and localized dispersal. *J. Evol. Biol.* 13, 814–825.
- Seeger, J., 1981. Kinship and covariance. *J. Theor. Biol.* 91, 191–213.
- Taylor, P.D., 1992a. Altruism in viscous populations: an inclusive fitness model. *Evol. Ecol.* 6, 352–356.
- Taylor, P.D., 1992b. Inclusive fitness in a homogeneous environment. *Proc. R. Soc. Ser. B* 249, 299–302.
- Taylor, P.D., Irwin, A., 2000. Overlapping generations can promote altruistic behavior. *Evolution* 54, 1135–1141.
- Taylor, P.D., Irwin, A., Day, T., 2000. Inclusive fitness in finite demographically structured and stepping-stone populations. *Selection* 1, 83–93.
- Taylor, P.D., Day, T., Wild, G., 2007a. Evolution of cooperation in a finite homogeneous graph. *Nature* 447, 469–472.
- Taylor, P.D., Day, T., Wild, G., 2007b. From inclusive fitness to fixation probability in homogeneous structured populations. *J. Theor. Biol.* 249, 101–110.
- Wild, G., Taylor, P.D., 2004. Fitness and evolutionary stability in game theoretic models of finite populations. *Proc. R. Soc. Ser. B* 271, 2345–2349.
- Wilson, D.S., Pollock, G., Dugatkin, L.A., 1992. Can altruism evolve in purely viscous populations? *Evol. Ecol.* 6, 341–351.
- Wolfram Research, Inc., 2005. *Mathematica*, Version 5.2. Champaign, IL.