AMERICAN MATHEMATICAL SOCIETY

BULLETIN OF THE AMS
CURRENT EVENTS

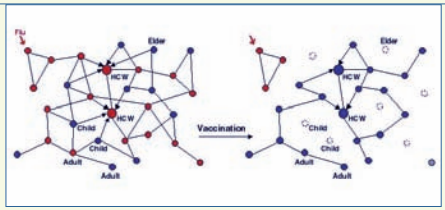
Saturday January 14, 2006, 1:00 p.m. – 5:50 p.m.

Organized by *David Eisenbud,* Mathematical Sciences Research Institute

1:00 PM

Lauren Ancel Meyers

Contact network epidemiology: Bond percolation applied to infectious disease prediction and control



2:00 PM

Kannan Soundararajan

Small gaps between prime numbers

3:00 PM

Madhu Sudan

Probabilistically checkable proofs

4:00 PM

Martin Golubitsky

Symmetry in neuroscience

Contact network epidemiology:
Bond percolation applied to infectious disease prediction and control.

Lauren Ancel Meyers

Abstract. Mathematics has long been an important tool in infectious disease epidemiology. I will provide a brief overview of compartmental models, the dominant framework for modeling disease transmission, and then contact network epidemiology, a more powerful approach that applies bond percolation on random graphs to model the spread of infectious disease through heterogeneous populations. I will derive important epidemiological quantities using this approach and provide examples of its application to issues of public health.

Background

Infectious diseases can have devastating impacts on human life and welfare. In the last three years, SARS, avian influenza, simian foamy virus, and monkeypox have jumped from animals into human populations. The uneven spread of SARS worldwide poignantly demonstrated that containment is possible, but depends critically on appropriate and aggressive management. With the growing threats of newly emerging diseases and bioterrorism, strategies to rapidly and effectively control outbreaks are vital to public health.

Mathematics is an invaluable epidemiological tool. It allows public health officials to conduct virtual experiments that would be practically unfeasible or unethical. Controlled experiments to evaluate the efficacy of control strategies are impossible in practice as we cannot intentionally introduce disease into populations or withhold potentially lifesaving interventions for the sake of scientific study. Mathematical models of disease transmission dynamics enable systematic evaluation of strategies such as vaccination and quarantine, and thereby provide a way around this difficulty.

In the 18th century, Daniel Bernoulli – the son, nephew and brother of mathematicians Johann, Jacob and Nicolaus II Bernoulli, respectively – made one of the first great mathematical contributions to infectious disease control [1]. While formally trained in medicine, Bernoulli is known for his research in biomechanics, hydrodynamics, economics, and astronomy. He also played an important role in the eradication of smallpox from Europe, which was likely introduced there in the early 16th century, and was endemic (maintained constantly) by the 18th century. Variolation is an inoculation technique whereby a scab or pus from an individual with a mild smallpox infection is introduced into the nose or mouth of healthy individuals. This practice began as early as 1000 AD in China and India and was introduced into England in 1717, where it was initially controversial. While variolation reduced the mortality probability of infected individuals from 30% to 1% [2], there was a small chance that the procedure would lead to death from a full-blown case of smallpox.

Bernoulli developed a mathematical model with which he argued that the gain from variolation in life expectancy through the eradication of smallpox far outweighed associated risks [1, 3]. Assuming that all individuals had a one in n chance of catching smallpox, and a one in m chance of dying from an infection, he derived the following

equation for the change in the number of currently naïve (never been infected) individuals in a specific age cohort during a small increment of time:

$$-\mathrm{d}s = \frac{s\mathrm{d}x}{\mathrm{n}} + \frac{s}{\xi}\left(\left(-\mathrm{d}\xi\right) - \frac{s\mathrm{d}x}{mn}\right) \tag{1}$$

where dx is the change in the age of the individuals in the cohort, s and ds are the number of currently naïve individuals and the change in that number, respectively, and ξ and $\mathrm{d}\xi$ are the total number of individuals in the cohort and the change in that number, respectively. On the right side of the equation, the first term is the number of new infections and the second term is the loss of susceptible individuals through death from other causes. Bernoulli integrated equation (1) and assumed that each cohort is born entirely susceptible (that is, $s = \xi$ when x=0) to find the expected fraction of susceptible individuals in a cohort of age x. This fraction is given by

$$\frac{s}{\xi} = \frac{m}{(m-1)\mathrm{e}^{\frac{x}{n}} + 1}. \tag{2}$$

Bernoulli assumed that the risk of catching smallpox was 12.5% (one in eight) in a given year across all age classes and that the mortality rate was 12.5% (one in eight) for all infected individuals. Using overall survivorship estimates calculated by Edmund Halley (of comet fame), he then used equation (2) to predict the mortality rates in every age class in a steady-state population with a birth class of size 1300.

Inoculation via variolation of all newborns would confer widespread immunity, yet entail some mortality due to variolation-induced smallpox. Bernoulli compared the annual mortality rates and average life expectancy predicted by his model to those predicted assuming universal inoculation, and found that variolation saves lives even if the mortality rate associated with variolation is quite high (with his parameters, as high as 10.6%).

Bernoulli's calculations clarified the benefits of widespread inoculation, even when there are significant risks. England began widely administering variolation in the 1750's, and upon the development of the smallpox vaccine in 1796, mandated the inoculation of all infants. Thanks to these efforts, smallpox was eradicated by the end of the 19th century.

Since Bernoulli, mathematicians and statisticians have offered many practical insights into infectious disease control. Notably, the English statistician William Farr analyzed the spatial distribution of choera cases, and thereby provided the first solid evidence that the disease spread via water rather than air [4].

Mathematical epidemiology exploded in the 20th century following the introduction of an intuitive and tractable framework. Between 1906 and 1927, the mass-action principle was introduced [5] and ultimately formalized in a deterministic model of disease transmission now attributed to Kermack and McKendrick [6]. In chemistry, the mass-action *law* states that the rate of a chemical reaction is proportional to the product of the concentrations of the reacting substances. In epidemiology, the mass-action *assumption* states that the number of new cases of disease in a time interval is proportional to the product of numbers of infected and susceptible hosts in the previous time interval. Within a decade, Reed and Frost introduced the first stochastic version of this model, the chain-binomial, which assumes that a disease spreads in discrete generations [7, 8]. This model derives a probability law for the next generation from that

of the present generation. More recently, Anderson and May among others have extended these efforts into a flexible approach, known as compartmental modeling, for predicting the transmission of a wide range of diseases on multiple scales [9].

We will discuss this framework as well as some of its practical applications and limitations below. This will set the stage for the introduction of contact network epidemiology, a new analytical approach that overcomes a major limitation of the mass-action assumption.

Compartmental SIR models

Compartmental models subdivide host populations by disease status. A simple and widely used example is the *SIR* model that tracks the movement of hosts among three states: susceptible (S), infected (I), and recovered (R) (Figure 1) [6]. These models assume that upon infection, hosts are immediately infectious and remain infectious until they recover. Infected hosts are assumed to have potentially disease-causing contacts with *random* individuals from the population according to a Poisson process that yields an average contact rate of β per unit time. Disease transmission occurs if and only if the individual at the receiving end of the contact is susceptible. There lies the mass-action assumption.

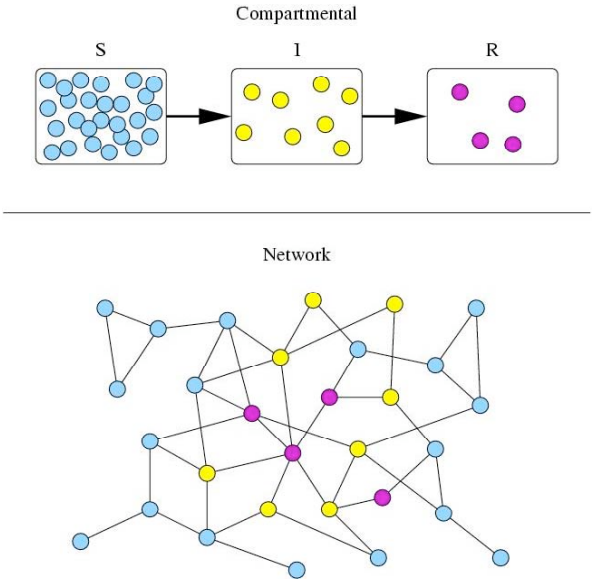


Figure 1. Compartmental and contact network models. Mass-action models assume that all individuals in a group are equally likely to become infected, while contact network epidemiology considers diverse contact patterns that underlie disease transmission. The disease spreads along the arrows (top) and the edges (bottom). (S=susceptible, I=infected, and R=recovered.)

Infectious hosts leave the infectious state at an average rate ν either by recovering and becoming immune or by dying. Thus the recovered class is a catchall for hosts that have been previously infected and are no longer infected or susceptible. In the limit of a large host population, this process is modeled by the following coupled nonlinear differential equations:

$$\begin{aligned} \frac{dS}{dt} &= -\beta IS, \\ \frac{dI}{dt} &= \beta IS - \nu I, \\ \frac{dR}{dt} &= \nu I, \end{aligned} \tag{3}$$

where $S(t)$, $I(t)$, and $R(t)$, are the numbers of susceptible, infected, and recovered hosts, respectively. Because the model ignores the birth and death of susceptibles, the total population size $N = S + I + R$ is static, and therefore the third equation is unnecessary. These equations apply to rapidly spreading diseases like measles and influenza that confer immunity extending beyond the typical length of an epidemic. The model can be easily adapted to consider the loss of immunity as well as birth and death dynamics.

The Basic Reproductive Rate

One of the touchstones of epidemiology is the *basic reproductive rate* of a disease: the number of secondary infections produced by a single infected host in an entirely susceptible population. This quantity indicates the initial growth rate for the infected class and the potential for a large-scale epidemic. In model (3), the per capita increase of infected individuals is given by

$$\frac{1}{I}\frac{dI}{dt} = \beta S - \nu. \tag{4}$$

The number of infected individuals increases by the product of the disease-causing contact rate β and number of susceptibles S and decreases by the combined recovery and mortality rate (henceforth *removal rate*) ν, which has units of $1/t$. The reciprocal of the removal rate, $1/\nu$, is the average time interval during which an infected individual remains contagious. The number of secondary cases infected per unit time is βS which yields a basic reproductive rate of

$$R_0 = \frac{\beta S}{\nu}. \tag{5}$$

If $R_0 > 1$, then each infected host will transmit disease to at least one other host during the infectious period, and the model predicts that disease will spread through the population. If not, then the disease is expected to fizzle out before reaching a substantial fraction of the population. Thus $R_0 = 1$ is a critical epidemiological value. In other words, pathogens with high levels of contagion and low recovery and mortality rates will pose the greatest threat.

Herd Immunity

The immunization of a single host not only protects that host but also indirectly protects others against the possible of disease transmission from the immunized host. If a sufficient fraction of a population is immunized, then an epidemic may be averted altogether. The protection of an entire population via the immunity of a fraction of the population is called herd immunity.

Equation (5) can be rearranged to find the minimum size of a susceptible population necessary for an epidemic to occur. Assuming that $R_0 = 1$, this threshold is given by

$$S_T = \frac{\nu}{\beta}. \tag{6}$$

A pathogen will go extinct if the size of the susceptible population is less than this threshold ($S < S_T$). If the population size is above this threshold, then we can rewrite the basic reproductive rate as

$$R_0 = \frac{S}{S_T}. \tag{7}$$

Immunization reduces the size of the susceptible class, and thus leads to a smaller basic reproductive rate of the pathogen. In particular, immunizing a fraction *p* of a population reduces R_0 to

$$R_0^i = \frac{(1-p)S}{S_T} = (1-p)R_0. \tag{8}$$

Immunization will successfully eradicate the disease if it causes the basic reproductive rate to drop below one. Thus the critical immunization rate p_c is

$$p_c = 1 - \frac{1}{R_0}. \tag{9}$$

Extensions of this basic model have been used to predict the minimum coverage necessary to drive specific diseases to extinction. For example, measles and whooping cough—two of the most contagious diseases—are thought to require 90-95% coverage, chicken pox and mumps 85-90% coverage, polio and scarlet fever 82-97% coverage, and smallpox 70-80% coverage [9].

Limitations of the Mass-Action Assumption: The example of SARS

Shortly after severe acute respiratory syndrome (SARS) was first recognized outside of Asia, mathematical epidemiologists estimated the average number of secondary cases emanating from one primary case in a susceptible population (R_0) to be in the range of 2.2 and 3.6 for this virus – an estimate well above one, approximating that of a new subtype of influenza [10-12]. Despite this estimate and near-universal susceptibility, SARS did not emerged as a global pandemic. Instead, initial seeding was followed by intense but tightly circumscribed activity in some locales with only scant activity in others.

The discrepancy between the estimates of R_0 and the observed epidemiology might stem from early and effective intervention since R_t, the reproductive ratio of a disease at time *t*, will decrease with the implementation of successful infection control measures. Yet, even during the three and a half months in which SARS spread in China

between its initial appearance and the broad implementation of public health measures, case counts were much less than expected from such values of R_0 [13], as shown by a back-of-the-envelope calculation. By definition, the total number of expected cases of a disease goes up by a factor of R_0 for every generation of infection, a generation being the mean time between an individual becoming infected and their infecting others. Based on recorded dates of the first symptoms for 124 pairs of subsequent infections in Singapore and Toronto [14, 15], the average generation time (γ) for SARS is estimated to be 9.7±0.3 days. This estimate clearly depends on the accuracy of the reported data. Roughly, the cumulative number of SARS cases over D days should be

$$\sum_{i=0}^{D/\gamma} (R_0)^i = \frac{1 - R_0^{D/\gamma + 1}}{1 - R_0}. \qquad (10)$$

(This is capped by total population size and does not consider the reduction in R_t once a substantial proportion of the population is infected). Thus for R_0 ranging between 2.2 and 3.6, this equation predicts that in the first 120 days of transmission in China, there should have been between approximately 30,000 and 10 million cases. In fact only 782 cases were reported during the initial three months [16], which, using this simple calculation, suggests that R_0 should be much lower and closer to 1.6.

Why do the initial estimates of R_0 seem incompatible the observed epidemiology in China? The basic reproductive rate has two basic inputs: (1) intrinsic properties of the pathogen that determine the transmission efficiency per contact and the duration of the infectious period and (2) the patterns of contacts between infected and susceptible hosts in the population. While the first factor may be fairly uniform across outbreaks, the second may be quite context dependent, varying both within and among populations. The problem with the SARS estimates stems from the mass-action assumption of compartmental models – that all individuals in a group are equally likely to become infected (or infect others) – often does not hold and therefore may lead to spurious estimates or estimates that cannot justifiably be extrapolated from the specific setting in which they were measured to the broader community context. Early SARS estimates were based largely on transmission data from closed settings like hospitals and crowded apartment buildings, where there are unusually high rates of contact between individuals [10, 11]. In fact, hospital transmission accounts for 50% of the value of R_0 described in [11]. If the contact patterns within these settings vary considerably, then the estimates for R_0 may be inaccurate. Even if the estimates for R_0 were indeed appropriate for these specific settings, they probably should not be extrapolated to the population at large. Contact rates may be considerably lower outside hospitals and crowded apartment buildings and, thus, so may be the general value of R_0 for SARS [17].

SARS, like many other infectious diseases, exhibited great heterogeneity in transmission efficiency with certain individuals appearing to be responsible for a large proportion of transmission events [14, 18, 19]. These individuals may be "superspreaders" with unusually large numbers of contacts or "supershedders" who are unusually effective at excreting the virus into the environment they share with others. In contrast to the mass-action assumption of standard compartmental models, the contact patterns in a community may be quite diverse. There is an enormous difference between a situation in all individuals share typical contact patterns and one in which most infected individuals pass the disease on to only one or even zero others, but a small number pass it

onto dozens or even hundreds – the mean value of R_0 can be the same in both cases, while the epidemiological outcomes are vastly different.

While the mass-action assumption laid the groundwork for major advances in epidemiological theory, it may be inappropriate when contact patterns are heterogeneous. To overcome this limitation, mathematical epidemiologists have developed several methods to explicitly consider heterogeneity in contact patterns. To name a few, more complex deterministic and stochastic compartmental models with multiple demographic groups capture greater contact heterogeneity [8, 20], other stochastic approaches including branching process models [21, 22], dyad models [23, 24], and Reed-Frost chain-binomial models [25] allow better predictions of the size and probability of epidemics; and "individual-based modeling", a computational approach which tracks the contact and infection histories of simulated individuals, yields detailed statistical predictions about disease outcomes [26-32]. Today we will consider a recent addition to this toolkit, *contact network epidemiology*, which is an analytical framework that explicitly and intuitively captures the diverse interactions that underlie the spread of diseases (Figure 1) [33-41].

Contact network epidemiology

The methods of contact network epidemiology can be divided into three steps. First we attempt to build a realistic network (graph) model of the contact patterns at an appropriate temporal and spatial scale. Second, we mathematically predict the spread of disease through the population based on intrinsic features of the pathogen and structural properties of the network. Third, we manipulate the network to model control strategies and analyze the epidemiological impact of such manipulations. We will now discuss each of these steps with illustrative examples.

The contact networks

A contact network model captures the patterns of interactions can lead to the transmission of an infectious disease. In a contact network, each person (or location) translates into a "vertex" and contacts among people (or locations) translate into "edges" that connect appropriate vertices. For example, one might model the contacts between individuals in a hospital or city that might lead to respiratory disease transmission [32, 41-43], the contacts between different geographical regions via human travel patterns that might lead to long-range transmission, or the sexual interactions within a high school that might lead to sexually transmitted disease transmission [44, 45].

The number of edges emanating from a vertex is called the *degree* of the vertex and indicates the number of possible contacts that can lead to disease transmission to or from an individual. The distribution of the number of such contacts within a population (the *degree distribution*) is fundamental to the ability of disease to spread through the population. The mass-action assumption of compartmental models is tantamount to assuming that the underlying contact patterns form a random graph with a Poisson degree distribution. If a network departs significantly from this ideal structure, then the traditional modeling approach may be invalid.

The contact (or social) network is a hot concept across many disciplines including sociology, epidemiology, biology, computer science and physics [46]. Researchers look for universal properties, and have paid special attention to small-world networks—

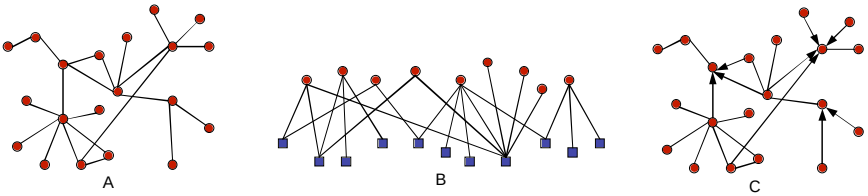


Figure 2. Contact networks. (A) Undirected network; (B) Bipartite network; and (C) Semi-directed network.

characterized by high levels of both local clustering and global connectivity [47], and scale free networks—characterized by degree distributions that follow a power law distribution with a small fraction of very highly connected hubs [48]. Several epidemiological-relevant contact networks including sexual contact networks and the internet, for example, have been characterized as scale free [49-51].

Realistic contact networks, however, do not always fall into one of these well-studied families of networks [41, 43]. Some have more complex structures, for example, those depicted in Figure 2. Bipartite networks, in which there are two types of nodes, have been used to represent asymmetric probabilities of transmission between caregivers and patients in a medical facility [43]. Semi-directed networks, in which some contacts are reciprocal and others are unidirectional, have been used to capture situations in which a person may infect another person but the converse is not true [42]. This situation may arise, for example, when infected individuals seek medical treatment during an outbreak. Suppose individual A is normally healthy and thus has no reason to go to the hospital until he or she becomes infected. At that point, individual A may come into contact and potentially spread disease to caregivers at the hospital. In contrast, if a caregiver at the hospital acquired the disease while individual A remained healthy, then there would be no opportunity for transmission in the opposite direction. This asymmetry can be modeled by directed edges pointing from individual A to health care workers. As described next, the mathematical methods of contact network epidemiology can accommodate such complex random networks with arbitrary degree distributions.

Predicting disease dynamics

Imagine that an infectious disease first appears at a randomly chosen vertex in a contact network (epidemiologically speaking, that vertex represents *patient zero*). Disease will propagate through the network as described for the compartmental models, except that the Poisson distribution of contacts is replaced by the structure of the contact network. The initial vertex will remain infected and infectious for some period of time, during which it has the potential to transmit disease to each of its contacts. The secondary cases likewise can transmit disease to their contacts during their infectious period, and so on. This process resembles bond percolation and can be analyzed using percolation models from statistical physics [35, 41-43, 52, 53]. This approach was initially suggested by Grassberger [54], and has been extended into a flexible framework for infectious disease prediction by Newman and colleagues [42, 43]. In what follows, we will review

Newman's derivations of fundamental epidemiological quantities for an undirected random network with an arbitrary degree distribution and some practical corollaries.

The percolation of disease through a network depends on both the level of contagion and the structure of the contact network. Following Newman [35], every edge in a network has a per unit time probability of disease transmission associated with it (r_{ij}), that is, the probability that vertex *i*, if infected, will transmit disease to vertex *j* in a given time increment. Assuming discrete time steps, if vertex *i* is infectious for τ time steps, then the probability that *j* will be infected by *i* is $T_{ij} = 1-\left(1-r_{ij}\right)^{\tau}$. For continuous time, $1-T_{ij} = \lim_{\delta t\to 0}\left(1-r_{ij}\delta t\right)^{r/\delta t} = \mathrm{e}^{-r_{ij}\tau}$, and thus $T_{ij} = 1-\mathrm{e}^{-r_{ij}\tau}$. The quantity r_{ij} summarizes core aspects of disease transmission including the likelihood that a contact will lead to transmission and individual susceptibility and will therefore vary across individuals. If r_{ij} is assumed to be an independent identically distributed (iid) random variable chosen from a distribution $P(r)$, then T_{ij} is also an iid random variable. Therefore the spread of disease will depend only on the mean probability of transmission between individuals (henceforth, the average transmissibility), which is given by

$$T = \left\langle T_{ij}\right\rangle = 1-\int_0^{\infty} Q(r)\,dr \tag{11}$$

where $Q(r) = 1-P(r)(1-r)^{\tau}$ or $Q(r) = 1-P(r)\mathrm{e}^{-r\tau}$, for discrete or continuous time, respectively.

Probability generating functions

To predict the fate of an outbreak, we use probability generating functions (pgf's), quantities that describe probability distributions, and here, summarize useful information about the structure of the contact network. The pgf for the degree distribution of a network is

$$G_0(x) = \sum_{k=1}^{\infty} p_k x^k \tag{12}$$

where p_k is the relative frequency of vertices of degree *k* in the network. The average degree equals the derivative of this function at *x*=1, that is, $\langle k\rangle = \sum_{k=1}^{\infty} kp_k$.

If we choose a random edge and follow it to one of its vertices, then the number of remaining edges connected to the vertex is called the *excess degree* of the vertex. The higher the degree of a vertex, the more likely it is to lie at the end of a randomly chosen edge. In particular, the likelihood of reaching a vertex with degree *k*, and thus with excess degree *k*-1 will be proportional to *k*. Thus the probability that a vertex at the end of a random edge has excess degree *k*-1 is $\frac{kp_k}{\langle k\rangle}$. This yields a generating function for the excess degree of a vertex of

$$G_1(x) = \frac{\sum_{k=1}^{\infty} kp_k x^{k-1}}{\sum_{k=1}^{\infty} kp_k}. \tag{13}$$

and an average excess degree of $\langle k_e \rangle = \frac{\sum_{k=1}^{\infty} k(k-1)p_k}{\sum_{k=1}^{\infty} kp_k} = \frac{\langle k^2 \rangle}{\langle k \rangle} - 1$.

When disease is introduced into a network, it will traverse some but not all of the edges according to the average transmissibility *T*. The edges that are infected during an epidemic are called *occupied*. Once the disease has run its course, the cluster of vertices connected to the first infected vertex by a continuous chain of occupied edges is exactly the outbreak. Ultimately we will characterize the size and distrubtion of this occupied cluster. We begin by deriving the pgf for the distribution of occupied edges attached to a randomly chosen vertex as a function of the average transmissibility *T*. The probability that a vertex has *m* of its *k* edges occupied is simply $\binom{k}{m} T^m (1-T)^{k-m}$, which leads to a probability generating function for the occupied degree of a vertex of

$$G_0(x;T) = \sum_{m=0}^{\infty} \sum_{k=m}^{\infty} p_k \binom{k}{m} T^m (1-T)^{k-m} x^m = G_0(1+(x-1)T). \tag{14}$$

Analogously, the pgf for the excess occupied degree, that is, the number of occupied edges emanating from a vertex reached by following a randomly chosen edge is given by

$$G_1(x;T) = G_1(1+(x-1)T). \tag{15}$$

Predicting the fate of a small outbreak

In general, percolation theory describes the behavior of connected groups of vertices in a random graph, and thus can be applied to predict the size of the infected cluster, that is, the number of vertices reached via disease transmission along the edges in the network. For a fixed network, there typically exists a threshold transmission rate below which only small, finite-sized outbreaks occur and above which large-scale epidemics (comparable to the size of the entire network) are possible.

First we will consider Newman's derivation of the epidemic threshold and the expected size of small outbreaks below the threshold. These calculations assume that mildly-contagious diseases spread in a tree-like fashion, causing only short transmission chains that do not loop back on themselves. Later, we relax this assumption and turn to diseases that lie above the epidemic threshold.

Let *s* denote the number of vertices contained in a small outbreak that begins at a randomly selected vertex and let $H_0(x;T)$ be the generating function for the distribution of outbreak sizes. Then

$$H_0(x;T) = \sum_s P_s(T) x^s \tag{16}$$

where $P_s(T)$ is probability that a single initial case sparks an outbreak of size *s* at the specified average transmissibility *T*. Let $H_1(x;T)$ be the generating function for the size of the cluster of connected vertices at the end of a randomly chosen edge.

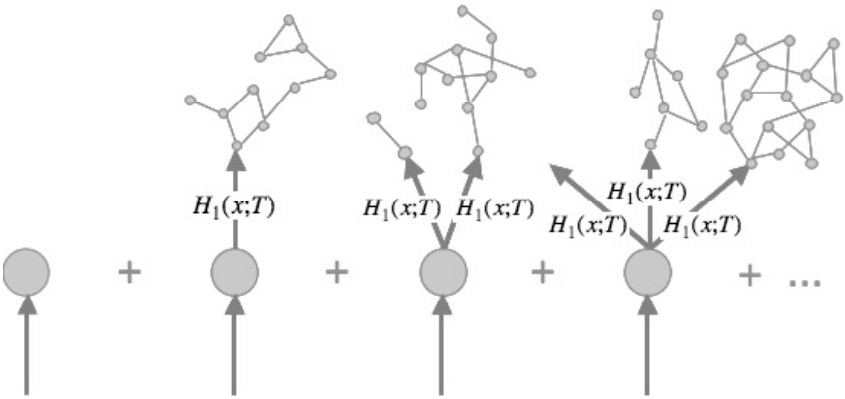


Figure 3. Future transmission diagram. When disease reaches an edge, we can consider all possible patterns of future transmission. The disease may not spread along the edge in the first place, it may spread along the edge but no further, it may spread along the original edge and then subsequently along another edge, it may spread along the original edge and then subsequently along two edges, etc. We can construct recursive equations to consider all possible outcomes.

To solve for the average value of *s*, we consider an outbreak that originates with a transmission event along a randomly chosen edge. The set of vertices reached by occupied edges can be represented in graphical form as in Figure 3. There are many possible outcomes: the disease does not spread along the edge, it spreads along the edge but no further, or it spreads along the edge and then subsequently along one or more additional edges emanating from the destination vertex. This is captured in a recursive equation

$$H_1(x;T) = xG_1\left(H_1(x;T);T\right). \tag{17}$$

This is roughly interpreted to mean that the size of a cluster proceeding from a randomly chosen edge *E* is the equal to sum of the sizes of the clusters at the end of each occupied edge emanating from the vertex *V* at the end *E* plus one for the vertex *V* itself. Likewise, the cluster emanating from a random vertex is generated by

$$H_0(x;T) = xG_0\left(H_1(x;T);T\right). \tag{18}$$

Consider now the average size $\langle s \rangle$ of an outbreak starting from a random vertex, which is given by

$$\begin{aligned} \langle s \rangle &= \sum_s sP_s(T) = H_0'(1;T) = 1 + G_0'(1;T)H_1'(1;T) \\ &= 1 + \frac{TG_0'(1)}{1 - TG_1'(1)} = 1 + \frac{T\langle k \rangle}{1 - T\left(\langle k^2 \rangle / \langle k \rangle - 1\right)} \end{aligned} \tag{19}$$

where the prime denotes differentiation with respect to the first variable. Note that for any normalized generating function $f(x)$, $f(1) = 1$. The expression for $\langle s \rangle$ diverges when the denominator in Equation (19) is zero, and only predicts the expected size of the outbreak when the denominator is greater than zero. Thus the equation

$$TG_1'(1) = 1 \tag{20}$$

marks the phase transition at which the size of an outbreak first becomes extensive. Hence an epidemic is only possible when the average transmissibility of a disease is greater than the critical transmissibility

$$T_c = \frac{1}{G_1'(1)} = \frac{\sum_k kp_k}{\sum_k k(k-1)p_k} = \frac{\langle k \rangle}{\langle k^2 \rangle - \langle k \rangle}. \tag{21}$$

We call T_c the *epidemic threshold*.

The basic reproductive rate

Recall that the basic reproductive rate is the number of secondary infections caused by a single infected host in a completely naïve population. In the contact network framework, this is simply the average number of occupied edges emanating from a vertex, that is,

$$R_0 = G_1'(1;T) = TG_1'(1) = T\left(\frac{\langle k^2 \rangle - \langle k \rangle}{\langle k \rangle}\right). \tag{22}$$

where $\langle k \rangle$ and $\langle k^2 \rangle$ are the mean degree and mean square degree (respectively) of the network. Recall that T_c is the critical transmissibility value above which population is vulnerable to large-scale epidemics (but is not guaranteed to experience an epidemic) and below which only small local outbreaks occur. If the transmissibility of a disease equals the epidemic threshold ($T=T_c$), then $R_0=1$.

Notice that the basic reproductive rate depends explicitly on the structure of the network (on $\langle k \rangle$ and $\langle k^2 \rangle$). A single pathogen may therefore have very different transmission dynamics depending on the population through which it spreads. If two networks have the same mean degree, $\langle k \rangle$, then the one with the larger variance in degree, $\langle k^2 \rangle - \langle k \rangle^2$, will be more vulnerable to the spread of disease. Estimates of R_0 that assume a mass-action model may therefore be invalid for populations with non-Poisson contact patterns, and in particular, underestimate the actual growth rate of the disease in highly heterogeneous networks.

Probability and size of a large-scale epidemic

When the transmissibility of a disease is larger than the epidemic threshold, then Equation (19) no longer indicates the size of the infected subpopulation. This is because transmission is so rampant that the chains of transmission are likely to loop back upon themselves, thus violating the assumption underlying the calculations depicted in Figure 4. When we are above the epidemic threshold, in the region in which epidemics can occur, we would like to know two quantities: the probability that a large-scale epidemic occurs and the fraction of individuals that are infected in that case. In an undirected network, these quantities are equal to each other and to the fraction of vertices from which an extensive numbers of others can be reached by following occupied edges. In the language of percolation, this is the giant component defined by occupied edges.

The probability of a full-blown epidemic, *S*, is derived by first calculating the likelihood that a single infection will lead to only a small outbreak instead of a full-blown

epidemic, and then subtracting that value from one. Recall that $H_0(x;T)$ is the generating function for the size of small outbreaks. Therefore $H_0(1;T)$ is the total probability that a randomly chosen initially infected vertex will lead to a finite sized outbreak. The probability of a large-scale epidemic is then given by

$$S = 1 - H_0(1;T) = 1 - G_0(u;T) \tag{23}$$

where $u = H_1(1;T)$. Thus u is the solution to the equation

$$u = G_1(u;T). \tag{24}$$

In terms of the degree distribution, the probability of a large-scale epidemic and the expected fraction of the network infected during such an epidemic is

$$S = 1 - \sum_k p_k \left(1 + (u-1)T\right)^k \tag{25}$$

where u is the solution to the self-consistency equation

$$u = \frac{\sum_k k p_k \left(1 + (u-1)T\right)^{k-1}}{\sum_k k p_k}. \tag{26}$$

We use numerical root finding methods to solve for u.

Other useful epidemiological quantities

We have recently extended Newman's results [35] to provide insight into other epidemiological processes. In particular, we have derived the probability of becoming infected and sparking an infection as a function of the degree of a vertex, the probability of an epidemic starting from an outbreak that is already underway, and the residual structure of a network after an epidemic has run its course.

The probability that an individual will spark an epidemic [41]. The probability ε_k that a patient zero with degree k will start an epidemic is equal to the probability that transmission of the disease along at least one of the edges emanating from the original vertex will lead to an epidemic. For any one of its k edges, $1-T$ is the probability that the disease does not get transmitted along the edge and Tu is the probability that even if disease is transmitted to the next vertex, it does not proceed into a full-blown epidemic. Thus

$$\varepsilon_k = 1 - (1 - T + Tu)^k. \tag{27}$$

The probability that a disease cluster will spark an epidemic [41]. The probability that an outbreak of size N will ignite an epidemic is $1 - \prod_{i=1}^{N} (1 - \varepsilon_{k_i})$ where k_i is the degree of individual i. This is just one minus the probability that none of the N infected individuals sparks an epidemic. If we know the number of current cases but not their contact patterns, then our best estimate for the probability of an epidemic is calculated similarly, with each of the $(1-\varepsilon_{k_i})$'s replaced with the probability that a typical infected individual does not spark an epidemic. The number of edges through which a typical infected individual can start an epidemic is given by the excess degree pgf, and the probability that one of those

edges will not give rise to an epidemic is $1-T+Tu$. Thus the probability that none of those edges will be a conduit to an epidemic is $\left(\frac{\sum_{k=1}^{\infty} kp_k(1-T+Tu)^{k-1}}{\sum_{k=1}^{\infty} kp_k}\right)$, and the probability that an outbreak of size N sparks an epidemic is $1-\left(\frac{\sum_{k=1}^{\infty} kp_k(1-T+Tu)^{k-1}}{\sum_{k=1}^{\infty} kp_k}\right)^N$.

Individual risk of infection [41]. The probability v_k that an individual with degree k will become infected during an epidemic is equal to one minus the probability that none of his or her k contacts will transmit the disease to him or her. The probability that a contact does not transmit the disease is equal to the probability $(1-u)(1-T)$ that the contact was infected but did not transmit the disease plus the probability u that the contact was not infected in the first place. Thus, a randomly chosen vertex of degree k will become infected with probability

$$v_k = \varepsilon_k = 1-(1-T+Tu)^k. \tag{28}$$

Frailty and interference [55]. If hosts are immunized following infection, an epidemic will change the structure of the *epidemiologically active network* (the remaining susceptible nodes and the edges that connect them). We characterize the structural evolution of a network due to an epidemic in terms of *frailty*—the degree to which highly connected individuals are more vulnerable to infection, and *interference*—the extent to which the epidemic cuts off connectivity among the susceptible population that remains following an epidemic. For a vertex that never becomes infected during an epidemic, we can distinguish between its original degree k and its degree in the residual network consisting of all nodes that remain uninfected by the epidemic, k_r. To understand the structural evolution of the network we derive two new network statistics: the mean *original* degree of individuals remaining in the residual network $\langle k \rangle_r$ and the mean *residual* degree of the individuals remaining in the residual network $\langle k_r \rangle_r$.

Recall that v_k is the probability that a randomly chosen vertex of degree k will become infected in an outbreak. The proportion of individuals with original degree k who remain in the residual network after an epidemic is given by $q_k = \frac{p_k(1-v_k)}{\sum_j p_j(1-v_j)}$, and thus the mean original degree in the residual network is given by

$$\langle k \rangle_r = \sum_k kq_k = \frac{\sum kp_k(1-v_k)}{\sum p_k(1-v_k)}. \tag{29}$$

To calculate the residual degree, we partition the original network into the vertices that are infected during the epidemic and those that remain uninfected, and then calculate the fraction of edges in the original network that begin and end in the uninfected set. Each edge in the network has two ends called stubs. Thus a node with degree k will have exactly k stubs attached to it, and the total number of stubs in the network is

$N\sum kp_k$ where N is the number of nodes in the network. For every one of the approximately $N\sum p_k v_k$ nodes in the infected partition, disease transmission to that node will necessarily have occurred along an edge with both an origin and destination stub (with the exception of the first infection. For simplicity, we ignore this exception.) Thus $N\sum p_k\left(k-2v_k\right)$ is the total number of stubs in the network excluding those that were a conduit for disease transmission during the epidemic. We refer to this quantity as the total number of uninfected stubs. The fraction of uninfected stubs attached to uninfected nodes is then

$$\frac{\sum kp_k(1-v_k)}{\sum p_k\left(k-2v_k\right)}.$$

Assuming that the network is randomly connected with respect to the original degree distribution, the fraction of edges that connect uninfected nodes to each other is this quantity squared. Thus, the average residual degree in the residual network is this fraction multiplied by the total number of stubs and divided by the number of nodes in the residual network,

$$\left\langle k_{\mathrm{r}}\right\rangle_{\mathrm{r}}=\left(\frac{\sum kp_k(1-v_k)}{\sum p_k(k-2v_k)}\right)^2\left(\frac{N\sum p_k(k-2v_k)}{N\sum p_k(1-v_k)}\right)=\frac{\left(\sum kp_k(1-v_k)\right)^2}{\sum p_k(k-2v_k)\sum p_k(1-v_k)}. \quad (30)$$

Finally, we define *frailty* to be the difference between the mean original degree in the original network and the mean original degree in the residual network, scaled by the mean original degree,

$$\phi=\frac{\langle k\rangle-\langle k\rangle_r}{\langle k\rangle}. \quad (31)$$

This parameter quantifies the extent to which high degree individuals are preferentially infected during an epidemic. We define *interference* to be the scaled difference between the mean original degree in the residual network and the mean residual degree in the residual network,

$$\theta=\frac{\langle k\rangle_r-\langle k_r\rangle_r}{\langle k\rangle}. \quad (32)$$

This quantity is the extent to which the epidemic has cut off connectivity among the remaining susceptible population.

Epidemiological dynamics on random networks. All of the quantities above pertain to the final outcome of an outbreak or epidemic. Erik Volz has recently developed a system of nonlinear differential equations to model the dynamical progression of a disease spreading through a random network with arbitrary degree distributions [56]. His model considers the state of each edge and each stub (one end of an edge) in the network. An edge is considered *occupied* if it has ever been a conduit for disease transmission, *refractory* if it is connected to a recovered vertex and is not occupied, and susceptible if it is neither occupied nor refractory. The state of stub depends on the state of its edge and on which end of the edge it occupies. The four equations of the model track the changing distribution of edge and stub states as disease percolates through the network. This model provides important insight into the interaction between population structure and

epidemiological dynamics and will be an important tool for optimizing the timing and targets of control measures.

Predictions on more complex contact networks

Newman calculated the basic epidemiological quantities for random networks in which transmission rate is correlated with the degree of either the infecting vertex or the vertex becoming infected [35]. We have subsequently made similar calculations for bipartite [43] and semi-directed contact networks [42]. Let us briefly consider some interesting features of disease transmission on semi-directed networks.

In a semi-directed network, each vertex has an *undirected degree* representing the number of undirected edges joining the vertex to other vertices as well as both an *in-degree* and an *out-degree* representing the number of directed edges incoming from other individuals and outgoing to other individuals respectively. The undirected-degree and in-degree indicate how many contacts can spread disease to the individual, and thus is related to the likelihood that an individual will become infected during an epidemic. The undirected-degree and out-degree indicate how many contacts may be infected by that individual should he or she become infected, and thus is related to the likelihood that an individual will contribute to an epidemic. The *semi-directed degree distribution* is the joint probability distribution p_{jkm} that a vertex has *j* incoming edges, *k* outgoing edges, and *m* undirected edges.

Our derivations in [42] reveal that semi-directed networks are more complicated than undirected networks in two important respects. First, there can be two different distributions of transmission rates—one for the directed edges and one for the undirected edges. When these distributions differ, the epidemic threshold is no longer a single value but a line dividing the two-dimensional space of transmission rates into a region in which there are only small outbreaks that die out before reaching a sizable fraction of the population and another region in which an epidemic is possible.

Second, recall that in an undirected network the probability of an epidemic and the expected fraction of the network infected during an epidemic are equal. In a semi-directed network, however, when the in-degree and out-degree distributions differ, then so do the probability of an epidemic and the expected incidence should one occur. These quantities are equivalent to the fraction of vertices from which an extensive numbers of others can be reached by following occupied edges and the fraction of vertices contained in such an extensive interconnected group, respectively. In the language of percolation, these are the giant strongly connected component (GSCC) plus the giant in-component (GIN) and the GSCC plus the giant out-component (GOUT) defined by occupied edges. Figure 4 illustrates the component structure of semi-directed networks. The relative size of the region shaded in vertical lines indicates the probability that any single infection will lead to a widespread epidemic, and the relative size of the region shaded in horizontal lines indicates the expected fraction of the population that will become infected during such an epidemic.

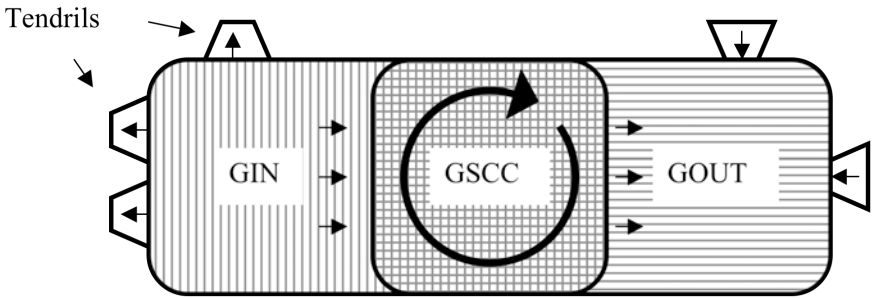


Figure 4. Structure of a semi-directed network. The largest set of vertices in which you can move between any two by following edges in the correct direction is the *giant strongly connected component* (GSCC). The set of vertices not contained in the GSCC that can be reached by following edges in the correct direction from the GSCC is called the *giant out-component* (GOUT). The set of vertices not contained in the GSCC from which the GSCC can be reached by following edges in the correct direction is called the *giant in-component* (GIN). Vertices that are not in the GSCC, GIN, or GOUT but can either be reached from the GIN or can reach the GOUT are in the *tendrils* of the network.

Evaluating control strategies

A primary public health goal is to bring disease from above an epidemic threshold value to below the threshold value, thereby eliminating the threat of a large-scale epidemic. This can be achieved through interventions that either directly impact the infectiousness of the pathogen, modify patterns of interaction so that the pathogen cannot easily spread through the population, or immunize segments of the population. We call these three forms of intervention *transmission reducing*, *contact reducing*, and *immunizing* [57].

Transmission reducing interventions introduce physical barriers to interrupt the spread of respiratory droplets or other infectious particles (e.g. face masks, gowns and gloves, hand hygeine, disinfection of animate objects). These interventions are modeled by reducing T_{ij} – the probability of transmission from vertex i to vertex j – at an appropriate subset of vertices.

Contact reducing interventions include isolation of infected persons, quarantine of persons during their incubation period, patient and/or staff cohorting in hospitals, closing public spaces (e.g. schools). These interventions are modeled by removing appropriate edges between vertices. For example, one can model school closures in an urban contact network by removing all edges that represent contacts between students, teachers, staff, etc. that take place during school.

Immunizing interventions include prophylactic medication and diverse vaccination strategies (e.g. ring vaccination – vaccinating individuals in contact with the identified infected case, targeted vaccination – vaccinating specific groups of individuals based on risk factors such as age, health, and place of employment, and general vaccination). Vaccination prior to an outbreak is tantamount to removing the immunized individuals from the network entirely, and thus is modeled by removing vertices corresponding to those individuals.

To evaluate candidate control measures, we first modify the contact network accordingly, and then quantify the impact of these changes on the size of an outbreak and demographic distribution of infections, identify segments of the population where compliance is most critical to successful control, and predict the individual and social benefits of complying with control measures. We will conclude with two practical examples.

Example 1: Controlling walking pneumonia outbreaks in closed settings
Walking pneumonia is a relatively mild form of pneumonia that spreads rapidly in closed settings such as hospitals, nursing homes, military communities, and college campuses. As with many diseases, conducting human experiments to test control measures is often infeasible or unethical. In collaboration with U.S. Centers for Disease Control and Prevention (CDC) officials we built some of the first network-based models of health-care settings with which we evaluated candidate strategies [43]. This work led the CDC to reject prior candidate strategies (including isolation of sick patients and antibiotic prophylaxis) in favor of the following intuitive yet previously overlooked strategy: upon the first diagnosis of walking pneumonia, reassign caregivers so that they limit their interactions to fewer wards. Although intuitive in retrospect, this insight came directly from analyzing disease transmission in a realistic model of the underlying network of interactions.

Example 2: Optimal distribution of influenza vaccines
Pandemic influenza is characterized by wide geographic person-to-person spread of a novel strain toward which the population has no immunity. The three major pandemics of the 20th century, in 1918, 1957 and 1968, collectively caused at least 600,000 deaths in the US and over 40 million deaths worldwide. Between major pandemics, the US experiences seasonal outbreaks of interpandemic flu that kill over 30,000 people annually.

The threat of an avian influenza pandemic and the 2004-2005 influenza vaccine supply shortage in the United States has sparked a debate about optimal vaccination strategies to reduce the burden of morbidity and mortality caused by the influenza virus. During the 2004-2005 shortage, the CDC restricted influenza vaccination to those most at risk for hospitalization and death — healthy infants, elderly, and individuals with chronic illnesses. These demographics, however, are not the primary spreaders of the influenza virus. Influenza outbreaks hinge, instead, on transmission by healthy school children [58-61], college students, and employed adults who may have many daily contacts and are more mobile [62]. Thus epidemiologists have suggested an alternative approach: vaccinate school-age children to slow the spread of the disease and thereby indirectly decrease mortality [63-65]. Several empirical studies support this strategy [66, 67]. Recently, Longini et al. used mathematical models to show that vaccinating 80% of all school-age children is almost as effective as vaccinating 80% of the entire population [63]. School-based vaccination programs have the additional benefits of high coverage, high efficacy and minimal side effects [68]. In a similar spirit, others have suggested contact-based priorities that target individuals with the highest numbers of potentially disease causing contacts [69-71], although this strategy may be difficult to implement.

Using an urban contact network based on demographic data for the city of Vancouver (with 260,000 individuals in 100,000 households), we have quantitatively

compared four strategies for both interpandemic and pandemic influenza [72]: (1) a mortality-based strategy that targets demographics with highest mortality rates (infants, elderly, and health care workers for interpandemic flu; and infants, adults, and health-care workers for pandemic flu); (2) a morbidity-based strategy, similar to the priorities suggested by Longini and Halloran [64] and Monto et al. [66], that targets school-aged children; (3) a mixed strategy that targets demographics with high attack rates (children) and high mortality rates (infants and elderly for interpandemic flu; infants and adults for pandemic flu); and (4) an idealized strategy that removes a fraction of the most connected individuals. For each of these strategies, we model immunization of 13% of the total population, based on reported coverage and efficacy levels for the targeted demographics [68, 73].

In contrast to prior studies [63], this study considers a relatively large population and the entire spectrum of viral transmission rates estimated for various influenza strains. As illustrated in Figure 5, the optimal strategy appears to depend critically on the viral transmissibility (reproductive rate) of the virus, with morbidity-based strategies outperforming mortality-based and mixed strategies for moderately transmissible strains, while the reverse is true for highly transmissible strains. This result holds for both interpandemic flu and pandemic flu. Furthermore, delays in vaccination and multiple introductions of disease into the community decrease the relative effectiveness of morbidity-based strategies. Thus, mortality-based strategies may be the prudent choice for outbreaks of new or atypical strains of influenza, when public health officials may not have reliable estimates for all (or any) of the first three inputs, and vaccination may be delayed. When reliable estimates of the key inputs are available significantly prior to an outbreak, then this approach can be applied to design optimal (rather than just prudent) priorities.

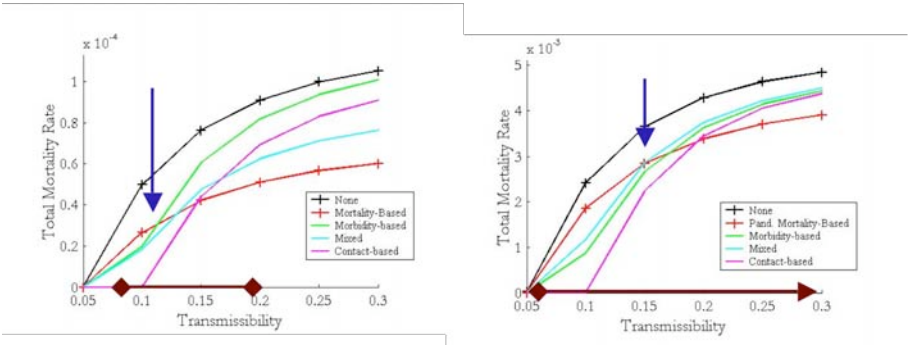


Figure 5. Expected mortality under different vaccination strategies for epidemic (left) and pandemic (right) influenza. Blue arrows mark transition between mildly-contagious diseases which are better controlled by morbitidy-based interventions and more highly contagious diseases which are better controlled by motality-based interventions. The brown arrows along the x-axis give the range of transmission rates estimated from influenza data from annual epidemics (left) and the 1918 pandemic (right).

Conclusions

Mathematical epidemiology continues to evolve, offering more detailed forecasting and more effective control. Much of the recent progress has been fueled by the importation of relatively simple ideas from dynamical systems, probability theory and statistical mechanics. Despite these steps forward, infectious disease control is more often that not, based on intuition rather than quantitative reasoning. This is particularly true for newly emerging diseases, for which we know little about the natural history and epidemiology of the pathogen. The variable public health response to SARS provides a compelling example of such uncertainty. When SARS emerged as a global threat in March 2003 [19], the WHO and other agencies issued travel warnings for affected cities. Hong Kong, Singapore and China closed schools [74]. A U.S. university denied attendance to students from China [75]. Public health authorities worldwide put thousands under quarantine [76]. Patients were strictly isolated [77] and specific hospitals in some cities were designated to receive SARS cases [78].

Contact network epidemiology can provide quantitatively grounded guidance for public health officials facing newly emerging diseases like SARS and avian influenza. Consider the previously cited analysis of flu vaccine strategies [72]. In contrast to other published mathematical approaches to this problem, our analytical methods have two advantages. They simultaneously capture realistic diversity in contact patterns ignored by many compartmental models and shortcut the extensive computer simulations required by agent-based models. In general, these advantages enable highly detailed and systematic consideration across several disease strains and intervention strategies. In the case of flu, such a study yielded important insight. If faced with a limited vaccine supply for either an interpandemic or pandemic strain of flu, morbidity-based strategies (e.g., targeting school children) are predicted to outperform mortality-based strategies (e.g., targeting elderly and infants) for strains that are mildly contagious, while the reverse is true for moderately to highly contagious strains. Furthermore, mortality-based strategies are generally advisable for populations experiencing repeated introductions of disease from other communities or delayed vaccination. This suggests that the US Centers for Disease Control's 2004 decision to prioritize the very young, the old, and the immunocompromised – those most at risk for complications from flu – is generally more prudent than the recently promoted alternative strategy of vaccinating school-children.

This methodology also sheds light on the incompatability between early estimates of R_0 for SARS and the case count in China (discussed above). This likely stemmed from the anomalously high contact rates in the hospital and apartment building upon which the R_0 estimates were based. Equation (22) clarifies that the basic reproductive rate of a disease is context dependent, that is, it fundamentally depends on the contact patterns of the population through which it spreads. Thus, while the SARS estimates may be valid for unusually crowded settings, they probably do not hold for typical rural or urban communities in general, like those through which SARS initially spread in China.

This example suggests that the emphasis on estimating the R_0 for an infectious disease may be misguided. Estimating the average transmissibility T instead of R_0 may be more valuable. This means reporting not just the number of new infections per case, but also the total estimated number of contacts during the infectious period of that case. Given the primary role of contact tracing in infectious disease control, the relevant data is often collected. Unlike R_0, T can be justifiably extrapolated from one location to another

even if the contact patterns are quite disparate. We offer a simple example to illustrate the benefits of measuring T. Suppose we measure R_0 =2.7 in a hospital where the average individual comes in close contact with 100 other individuals. Then the probability that an individual will catch the disease from an infected contact is just 2.7% or, in network terms, T=.027. Now suppose the typical individual in the general population has 10 close contacts that could potentially lead to the spread of a disease. If we extrapolate R_0=2.7 to the general public, then we predict that, on average, 2.7 out of every 10 contacts or 27% of contacts become infected. However, if we extrapolate T=.027 to the general public we still have only 2.7% of contacts becoming infected, which gives us a much reduced expectation for the spread of the disease.

In closing, mathematical epidemiology is a rapidly developing field that thrives on collaborations among scientists, mathematicians, and public health officials. Contact network epidemiology is a particularly promising approach in which progress is fueled by both scientific curiosity and public health concerns. As demonstrated by the variable response to SARS, there is need for greater quantitative reasoning in public health. The onus is the modelers not only to make technical advances, but also to demonstrate the utility and accessibility of our models.

References

1. Bernoulli, D. and S. Blower, An attempt at a new analysis of the mortality caused by smallpox and of the advantages of inoculation to prevent it. Reviews in Medical Virology, 2004. 14(5): p. 275-288.
2. Smallpox: A Great and Terrible Scourge. 2002 http://www.nlm.nih.gov/exhibition/smallpox/sp_variolation.html.
3. Dietz, K. and J.A.P. Heesterbeek, Daniel Bernoulli's epidemiological model revisited. Mathematical Biosciences, 2002. 180: p. 1-21.
4. Paneth, N. and P. Vinten-Johansen, A rivalry of foulness: Official and unofficial investigations of the London cholera epidemic of 1854. American Journal of Public Health, 1998. 88(10): p. 1545-1553.
5. Hamer, W.H., Epidemic disease in England-the evidence of variability and persistency of type. The Lancet, 1906. i: p. 733-739.
6. Kermack, W.O. and A.G. McKendrick, A contribution to the mathematical theory of epidemics. Proceedings of the Royal Society (London) A, 1927. 115: p. 700-721.
7. Abbey, H., An examination of the Reed-Frost Theory of Epidemics. Human Biology, 1952. 24(3): p. 201-233.
8. Bailey, N.T.J., The Mathematical Theory of Infectious Diseases and Its Applications. 1975, New York: Hafner Press.
9. Anderson, R.M. and R.M. May, Infectious Diseases of Humans, Dynamics and Control. 1991, Oxford: Oxford University Press.
10. Lipsitch, M., et al., Transmission Dynamics and Control of Severe Acute Respiratory Syndrome. Science, 2003: p. 1086616.
11. Riley, S., et al., Transmission Dynamics of the Etiological Agent of SARS in Hong Kong: Impact of Public Health Interventions. Science, 2003: p. 1086478.

12. Hethcote, H., Mathematics of infectious diseases. SIAM Review, 2000. 42: p. 599.
13. Xu, R.-H., et al., Epidemiologic Clues to SARS Origin in China. Emerg. Infect. Dis., 2004. 10(6).
14. Leo, Y.S., et al., Severe acute respiratory syndrome - Singapore, 2003. Morbidity and Mortality Weekly Report, 2003. 52(18): p. 405.
15. Poutanen, S.M., et al., Identification of Severe Acute Respiratory Syndrome in Canada. N. Engl. J. Med., 2003. 348(20): p. 1995.
16. World Health Organization. Severe Acute Respiratory Syndrome (SARS). 2003 http://www.who.int/csr/sars/en/.
17. Yu, I.T.S., et al., Evidence of airborne transmission of the severe acute respiratory syndrome virus. N. Engl. J. Med., 2004. 350(17): p. 1731-1739.
18. Booth, C.M., et al., Clinical Features and Short-term Outcomes of 144 Patients With SARS in the Greater Toronto Area. JAMA, 2003: p. 289.21.JOC30885.
19. Donnelly, C.A., et al., Epidemiological determinants of spread of causal agent of severe acute respiratory syndrome in Hong Kong. The Lancet, 2003: p. 1.
20. Hethcote, H.W. and J.A. Yorke, Gonorrhea Transmission Dynamics and Control. Lecture Notes in Biomathematics. Vol. 56. 1984, Berlin: Springer-Verlag.
21. Becker, N., Estimation for discrete time branching processes with application to epidemics. Biometrics, 1977. 33(3): p. 515-522.
22. Farrington, C., M. Kanaan, and N. Gay, Branching process models for surveillance of infectious diseases controlled by mass vaccination. Biostatistics, 2003. 4(2): p. 279-295.
23. Keeling, M.J., D.A. Rand, and A.J. Morris, Correlation Models for Childhood Diseases. Proceedings of the Royal Society (London) B, 1997. 264: p. 1149-1156.
24. Ferguson, N.M. and G.P. Garnett, More realistic models of sexually transmitted disease transmission dynamics: sexual partnership networks, pair models, and moment closure. Sex. Transm. Dis., 2000. 27(10): p. 600.
25. Lefevre, C. and P. Picard, On the formulation of discrete-time epidemic models. Mathematical Biosciences, 1989. 95(1): p. 27-35.
26. Kleczkowski, A. and B.T. Grenfell, Mean field-type equations for spread of epidemics: the 'small world' model. Physica A, 1999. 274(1-2): p. 1-385.
27. Durrett, R., Stochastic spatial models. SIAM Review, 1999. 41(4): p. 577-718.
28. Sander, L.M., et al., Percolation on heterogeneous networks as a model for epidemics. Mathematical Biosciences, 2002. 180: p. 293-305.
29. Ritton, T. and P.D. O'Neill, Bayesian Inference for Stochastic Epidemics in Populations with Random Social Structure. Scandanavian Journal of Statistics, 2002. 29(3): p. 375-390.
30. Van der Ploeg, C.P.B., et al., A microsimulation model for decision support in STD control. Interfaces, 1998. 28: p. 84-100.
31. Chowell, G., et al., Scaling laws for the movement of people between locations in a large city. Physical Review E, 2003. 68: p. 066102.
32. Eubank, S., et al., Modelling disease outbreaks in realistic urban social networks. Nature, 2004. 429: p. 180-184.
33. Keeling, M.J., et al., Modelling vaccination strategies against foot-and-mouth disease. Nature, 2003. 421(6919): p. 136-42.

34. Diekmann, O., M.C.M. de Jong, and J.A.J. Metz, A deterministic epidemic model taking account of repeated contacts between the same individuals. J. Appl. Prob., 1998. 35: p. 448-462.
35. Newman, M.E.J., Spread of epidemic disease on networks. Physical Review E, 2002. 66(1): p. art. no.-016128.
36. Sattenspiel, L. and C.P. Simon, The spread and persistence of infectious diseases in structured populations. Math. Biosci., 1988. 90: p. 341.
37. Ball, F., D. Mollison, and G. Scalia-Tomba, Epidemics with two levels of mixing. The Annals of Applied Probability, 1997. 7: p. 46.
38. Morris, M., Data driven network models for the spread of disease, in Epidemic Models: their structure and relation to data, D. Mollison, Editor. 1995, Cambridge University Press: Cambridge. p. 302-322.
39. Longini, I.M., A mathematical model for predicting the geographic spread of new infectious agents. Math. Biosci., 1988. 90: p. 367.
40. Lloyd, A.L. and R.M. May, Epidemiology. How viruses spread among computers and people. Science, 2001. 292(5520): p. 1316.
41. Meyers, L.A., et al., Network theory and SARS: predicting outbreak diversity. Journal of Theoretical Biology, 2005. 232: p. 71-81.
42. Meyers, L.A., M.E.J. Newman, and B. Pourbohloul, Predicting epidemics on directed contact networks. Journal of Theoretical Biology, in press. doi:10.1016/j.jtbi.2005.10.004.
43. Meyers, L.A., et al., Applying network theory to epidemics: Control measures for Mycoplasma pneumoniae outbreaks. Emerging Infectious Diseases, 2003. 9(2): p. 204.
44. Rothenberg, R.B., How a net works: implications of network structure for the persistence and control of sexually transmitted diseases and HIV. Sexually Transmitted Diseases, 2001. 28: p. 63-68.
45. Rothenberg, R.B., et al., Using social network and ethnographic tools to evaluate syphilis transmission. Sexually Transmitted Diseases, 1997. 25(3): p. 154-160.
46. Amaral, L.A.N. and J. Ottino, Complex networks - augmenting the framework for the study of complex systems. Eur Phys J B, 2004. 38: p. 147-162.
47. Watts, D.J., Small Worlds: The Dynamics of Networks Between Order and Randomness. 1999, Princeton: Princeton University Press.
48. Barabasi, A.L. and R. Albert, Emergence of scaling in random networks. Science, 1999. 286(5439): p. 509.
49. Liljeros, F., et al., The web of human sexual contacts. Nature, 2001. 411: p. 907-908.
50. Liljeros, F., C.R. Edling, and L.A.N. Amaral, Sexual networks: implications for the transmission of sexually transmitted diseases. Microbes, 2003.
51. Pastor-Satorras, R. and A. Vespignani, Epidemic spreading in scale-free networks. Phys Rev Lett, 2001. 86(14): p. 3200.
52. Sander, L.M., et al., Percolation on heterogeneous networks as a model for epidemics. Mathematical Biosciences, 2002. 180: p. 293-305.
53. Pourbohloul, B., et al., A quantitative comparison of control strategies for respiratory-borne pathogens. Emerging Infectious Disease, In press.
54. Grassberger, P., Critical behavior of the general epidemic process and dynamical percolation. Mathematical Biosciences, 1983. 63(2): p. 157-172.

55. Ferrari, M., et al., Netword fraily and the geometry of herd immunity. in review.
56. Volz, E., SIR dynamics in structured populations with heterogeneous connectivity. Journal of Mathematical Biology, in press.
57. Pourbohloul, B., et al., Modeling control strategies of respiratory pathogens. Emerging Infectious Disease, 2005. 11(8): p. 1249-1256.
58. Longini, I.M., et al., Estimating household and community transmission parameters of influenza. American Journal of Epidmiology, 1982. 115: p. 736-751.
59. Fox, J.P., et al., Influenza virus infections in Seattle families, 1975-1979. American Journal of Epidmiology, 1982. 116: p. 212-227.
60. Jennings, L.C. and J.A.R. Miles, A study of acute respiratory disease in the community of Port Chalmers. Journal of Hygiene, 1978. 81: p. 67-75.
61. Taber, L.H., et al., Infection with influenza A/Victoria virus in Houston families, 1976. Journal of Hygiene, 1982. 86: p. 303-313.
62. Glezen, W.P., Emerging infections: pandemic influenza. Epidemiology Reviews, 1996. 18(1): p. 64-76.
63. Longini, I.M., et al., Containing pandemic influenza with antiviral agents. American Journal of Epidmiology, 2004. 159: p. 623-633.
64. Longini, I.M. and M.E. Halloran, Strategy for distribution of influenza vaccine to high-risk groups and children. American Journal of Epidmiology, 2005. 161: p. 303-306.
65. Weycker, D., et al., Population-wide benefits of routine vaccination of children against influenza. Vaccine, 2005. 23(10): p. 1284-1293.
66. Monto, A.S., J.S. Koopman, and I.M. Longini, The Tecumseh study of illness. XIII. Influenza infection and disease, 1976-1981. American Journal of Epidmiology, 1985. 121: p. 811-822.
67. Reichart, T.A., et al., The Japanese experience with vaccinating school-children against influenza. New England Journal of Medicine, 2001. 344: p. 889-896.
68. Centers for Disease Control and Prevention, Prevention and control of influenza: recommendations of the Advisory Committee on Immunization Practices (ACIP). Morbidity and Mortality Weekly Report, 2004. 53: p. RR-6.
69. Dezso, Z. and A.L. Barabasi, Halting viruses in scale-free networks. Physical Review E, 2001. 65: p. 055103.
70. Pastor-Satorras, R. and A. Vespignani, Immunization of complex networks. Physical Review E, 2001. 65: p. 036104.
71. Albert, R., H. Jeong, and A.L. Barabasi, Attack and error tolerance of complex networks. Nature, 2000. 406: p. 378-382.
72. Bansal, S., B. Pourbohloul, and L.A. Meyers, Quantitative design of influenza vaccination programs. in review.
73. Centers for Disease Control and Prevention, Interim estimates of populations targeted for influenza vaccination from 2002 National Health Interview Survey Data and estimates for 2004 based on Influenza Vaccine Shortage Priority Groups. 2004.
74. SARS closing Beijing schools. 2003 http://www.cnn.com/2003/WORLD/asiapcf/east/04/22/sars.china.school/.
75. Berkeley turns away students fro SARS-hit regions. 2003 http://edition.cnn.com/2003/EDUCATION/05/05/berkeley.sars.ban.

76. CDC, Efficiency of quarantine during an epidemic of severe acute respiratory syndrome--Beijing, China, 2003. Morbidity and Mortality Weekly Report, 2003. 52(43): p. 1037-1040.
77. Public Health Guidance for Community-Level Preparedness and Response to Severe Acute Respiratory Syndrome (SARS) Version 2, Supplement D: Community Containment Measures, Including Non-Hospital Isolation and Quarantine. 2004 http://www.cdc.gov/ncidod/sars/guidance/D/pdf/lessons.pdf.
78. Designated hospitals in Beijing meeting SARS treatment demands. 2003 http://english.peopledaily.com.cn/200305/09/eng20030509_116407.shtml.

SMALL GAPS BETWEEN PRIME NUMBERS: THE WORK OF GOLDSTON-PINTZ-YILDIRIM

K. Soundararajan

Introduction. In early 2005, Dan Goldston, János Pintz, and Cem Yıldırım [12] made a spectacular breakthrough in the study of prime numbers. Resolving a long-standing open problem, they proved that there are infinitely many primes for which the gap to the next prime is as small as we want compared to the average gap between consecutive primes. Before their work, it was only known that there were infinitely many gaps which were about a quarter the size of the average gap. The new result may be viewed as a step towards the famous twin prime conjecture that there are infinitely many prime pairs p and $p+2$; the gap here being 2, the smallest possible gap between primes[1]. Perhaps most excitingly, their work reveals a connection between the distribution of primes in arithmetic progressions and small gaps between primes. Assuming certain (admittedly difficult) conjectures on the distribution of primes in arithmetic progressions, they are able to prove the existence of infinitely many prime pairs that differ by at most 16. The aim of this article is to explain some of the ideas involved in their work.

Let us begin by explaining the main question in a little more detail. The number of primes up to x, denoted by $\pi(x)$, is roughly $x/\log x$ for large values of x; this is the celebrated Prime Number Theorem[2]. Therefore, if we randomly choose an integer near x, then it has about a 1 in $\log x$ chance of being prime. In other words, as we look at primes around size x, the average gap between consecutive primes is about $\log x$. As x increases, the primes get sparser, and the gap between consecutive primes tends to increase. Here are some natural questions about these gaps between prime numbers. Do the gaps always remain roughly about size $\log x$, or do we sometimes get unexpectedly large gaps and sometimes surprisingly small gaps? Can we say something about the statistical distribution of these gaps? That is, can we quantify how often the gap is between, say, $\alpha \log x$ and $\beta \log x$, given $0 \le \alpha < \beta$? Except for the primes 2 and 3, clearly the gap between consecutive primes must be even. Does every even number occur infinitely often as a gap between consecutive primes? For example, the twin prime conjecture says that the gap 2 occurs infinitely. How frequently should we expect the occurrence of twin primes?

Number theorists believe they know the answers to all these questions, but cannot always prove that the answers are correct. Before discussing the answers let us address a

The author is partially supported by the National Science Foundation.

[1] apart from the gap between 2 and 3, of course!

[2] Here, and throughout, log stands for the natural logarithm.

possible meta-question. Problems like twin primes, and the Goldbach conjecture involve adding and subtracting primes. The reader may well wonder if such questions are natural, or just isolated curiosities. After all, shouldn't we be multiplying with primes rather than adding/subtracting them? There are several possible responses to this objection.

Firstly, many number theorists and mathematical physicists are interested in understanding spacing statistics of various sequences of numbers occuring in nature. Examples of such sequences are prime numbers, the ordinates of zeros of the Riemann zeta-function (see [21] and [23]), energy levels of large nuclei, the fractional parts of $\sqrt{n}$ for $n \le N$ (see [7]), etc. Do the spacings behave like the gaps between randomly chosen numbers, or do they follow more esoteric laws? Our questions on gaps between primes fit naturally into this framework.

Secondly, many additive questions on primes have applications to other problems in number theory. For example, consider primes p for which $2p+1$ is also a prime. Analogously to twin primes, it is conjectured that there are infinitely many such prime pairs p and $2p+1$. Sophie Germain came up with these pairs in her work on Fermat's last theorem. If there are infinitely many Sophie Germain pairs p and $2p+1$ with p lying in a prescribed arithmetic progression, then Artin's primitive root conjecture — every positive number a which is not a perfect square is a primitive root[3] for infinitely many primes — would follow. For example, if p lies in the progression $3 \pmod{40}$, and $2p+1$ is prime, then 10 is a primitive root modulo $2p+1$, and as Gauss noticed (and the reader can check) the decimal expansion of $1/(2p+1)$ has exactly $2p$ digits that repeat. There are also connections between additive questions on primes and zeros of the Riemann zeta and other related functions. Precise knowledge of the frequency with which prime pairs p and $p+2k$ occur (for an even number $2k$) has subtle implications for the distribution of spacings between ordinates of zeros of the Riemann zeta-function (see [1] and [23]). Conversely, weird (and unlikely) patterns in zeros of zeta-like functions would imply the existence of infinitely many twin primes (see [17])!

Finally, these 'additive' questions on primes are lots of fun, have led to much beautiful mathematics, and inspired many generations of number theorists!

Cramér's model. A useful way to think about statistical questions on prime numbers is the random — also known as Cramér — model. The principle, based on the fact that a number of size about n has a 1 in $\log n$ chance of being prime, is this:

The indicator function for the set of primes (that is, the function whose value at n is 1 or 0 depending on whether n is prime or not) behaves roughly like a sequence of independent, Bernoulli random variables $X(n)$ with parameters $1/\log n$ ($n \ge 3$). In other words, for $n \ge 3$, the random variable $X(n)$ takes the value 1 (n is 'prime') with probability $1/\log n$, and $X(n)$ takes the value 0 (n is 'composite') with probability $1 - 1/\log n$. For completeness, let us set $X(1) = 0$, and $X(2) = 1$.

This must be taken with a liberal dose of salt: a number is either prime or composite, probability does not enter the picture! Nevertheless, the Cramér model is very effective in predicting answers, although it does have its limitations (for example, if $n > 2$ is prime

[3]That is, a generates the multiplicative group of residues modulo that prime.

then certainly $n+1$ is not, so the events of n and $n+1$ being prime are clearly not independent) and sometimes leads to incorrect predictions.

Let us use the Cramér model to predict the probability that, given a large prime p, the next prime lies somewhere between $p+\alpha\log p$ and $p+\beta\log p$. In the Cramér model, let p be large and suppose that $X(p)=1$. What is the probability that $X(p+1)=X(p+2)=\ldots=X(p+h-1)=0$ and $X(p+h)=1$, for some integer h in the interval $[\alpha\log p,\beta\log p]$? We will find this by calculating the desired probability for a given h in that interval, and summing that answer over all such h. For a given h the probability we seek is

$$\Big(1-\frac{1}{\log(p+1)}\Big)\Big(1-\frac{1}{\log(p+2)}\Big)\cdots\Big(1-\frac{1}{\log(p+h-1)}\Big)\frac{1}{\log(p+h)}.$$

Since p is large, and h is small compared to p (it's only of size about $\log p$) we estimate that $\log(p+j)$ is very nearly $\log p$ for j between 1 and h. Therefore our probability above is approximately $(1-1/\log p)^{h-1}(1/\log p)$, and since $1-1/\log p$ is about $e^{-1/\log p}$, this is roughly

$$e^{-(h-1)/\log p}\Big(\frac{1}{\log p}\Big).$$

Summing over the appropriate h, we find that the random model prediction for the probability that the next prime larger than p lies in $[p+\alpha\log p,p+\beta\log p]$ is

$$\sum_{\alpha\log p\le h\le\beta\log p} e^{-(h-1)/\log p}\frac{1}{\log p}\approx\int_\alpha^\beta e^{-t}dt,$$

since the left hand side looks like a Riemann sum approximation to the integral.

Conjecture 1. *Given an interval $0\le\alpha<\beta$, as $x\to\infty$ we have*

$$\frac{1}{\pi(x)}\#\{p\le x: p_{\rm next}\in(p+\alpha\log p,p+\beta\log p)\}\to\int_\alpha^\beta e^{-t}dt,$$

where $p_{\rm next}$ denotes the next prime larger than p. Here, and throughout the paper, the letter p is reserved for primes.

We have deliberately left the integral unevaluated, to suggest that there is a probability density e^{-t} of finding $(p_{\rm next}-p)/\log p$ close to t. If we pick N random numbers uniformly and independently from the interval $[0,N]$, and arrange them in ascending order, then, almost surely, the consecutive spacings have the probability density e^{-t}. Thus, the Cramér model indicates that the gaps between consecutive primes are distributed like the gaps between about $x/\log x$ numbers chosen uniformly and independently from the interval $[0,x]$. In probability terminology, this is an example of what is known as a 'Poisson process.'

There are several related predictions we could make using the random model. For example, choose a random number n below x, and consider the interval $[n,n+\log n]$. The expected number of primes in such an interval is about 1, by the prime number theorem. But of course some intervals may contain no prime at all while others may contain several

primes. Given a non-negative number k, what is the probability that such an interval contains exactly k primes? The reader may enjoy the pleasant calculation which predicts that, for large x, the answer is nearly $\frac{1^k}{k!}e^{-1}$ — the answer is written so as to suggest a Poisson distribution with parameter 1.

Conjecture 1 makes clear that there is substantial variation in the gaps between consecutive primes. Given any large number Λ we expect that with probability about $e^{-\Lambda}$ (a tiny, but positive probability), the gap between consecutive primes is more than Λ times the average gap. Given any small positive number ϵ we expect that with probability about $1-e^{-\epsilon}$ (a small, but positive probability), the gap between consecutive primes is at most ϵ times the usual gap. Thus, two consequences of Conjecture 1 are

$$\limsup_{p\to\infty} \frac{p_{\text{next}}-p}{\log p} = \infty,$$

and

$$\liminf_{p\to\infty} \frac{p_{\text{next}}-p}{\log p} = 0.$$

Large gaps. Everyone knows how to construct arbitrarily long intervals of composite numbers: just look at $m!+2$, $m!+3$, ..., $m!+m$ for any natural number $m \geq 2$. This shows that $\limsup_{p\to\infty}(p_{\text{next}}-p)=\infty$. However, if we think of $m!$ being of size about x then a little calculation with Stirling's formula shows that m is about size $(\log x)/\log\log x$. We realize, with dismay, that the 'long' gap we have constructed is not even as large as the average gap of $\log x$ given by the prime number theorem. A better strategy is to take N to be the product of the primes that are at most m, and note again that $N+2$, ..., $N+m$ must all be composite. It can be shown that N is roughly of size e^m. Thus we have found a gap at least about $\log N$, which is better than before, but still not better than average. Can we modify the argument a little? In creating our string of $m-1$ consecutive composite numbers, we forced these numbers to be divisible by some prime below m. Can we somehow use primes larger than m to force $N+m+1$, $N+m+2$, etc., to be composite, and thus create longer chains of composite numbers? In the 1930s, in a series of papers Westzynthius [27], Erdős [8] and Rankin [25] found ingenious ways of making this idea work. The best estimate was obtained by Rankin, who proved that there exists a positive constant c such that for infinitely many primes p,

$$p_{\text{next}}-p > c\log p\frac{(\log\log p)\log\log\log\log p}{(\log\log\log p)^2}.$$

The fraction above does grow[4], and so

$$\limsup_{p\to\infty} \frac{p_{\text{next}}-p}{\log p} = \infty,$$

as desired. We should remark here that, although very interesting work has been done on improving the constant c above, Rankin's result provides the largest known gap between

[4]although so slowly that, as the joke goes, no one has observed it doing so!

primes. Erdős offered $10,000 for a similar conclusion involving a faster growing function. Bounty hunters may note that the largest Erdős prize that has been collected is $1,000, by Szemeredi [26] for his marvellous result on the existence of long arithmetic progressions in sets of positive density.

What should we conjecture for the longest gap between primes? Cramér's model suggests that

$$\limsup_{p\to\infty} \frac{p_{\text{next}} - p}{(\log p)^2} = c, \tag{1}$$

with $c = 1$. The rationale behind this is that the probability that $X(n) = 1$ and that the next 'prime' is bigger than $n + (1+\epsilon)\log^2 n$ is about $1/(n^{1+\epsilon}\log n)$, by a calculation similar to the one leading up to Conjecture 1. If ϵ is negative the sum of this probability over all n diverges, and the Borel-Cantelli lemma tells us that, almost surely, such long gaps occur infinitely often. If ϵ is positive, the corresponding sum converges and the Borel-Cantelli lemma says that almost surely we get these longer gaps only a finite number of times. More sophisticated analysis has however revealed that (1) is one of those questions which expose the limitations of the Cramér model. It appears unlikely that the value of c is 1 as predicted by the Cramér model, and that c should be at least $2e^{-\gamma} \approx 1.1229$ where γ is Euler's constant. No one has felt brave enough to suggest what the precise value of c should be! This is because (1) is far beyond what 'reasonable' conjectures such as the Riemann hypothesis would imply. An old conjecture says that there is always a prime between two consecutive squares. Even this lies (slightly) beyond the reach of the Riemann hypothesis, and all it would imply is that

$$\limsup_{p\to\infty} \frac{p_{\text{next}} - p}{\sqrt{p}} \le 4;$$

a statement much weaker than (1) with a finite value of c.

We cut short our discussion on long gaps here, since our focus will be on small gaps; for more information on these and related problems, we refer the reader to the excellent survey articles by Heath-Brown [18] and Granville [15].

Small gaps. Since the average spacing between p and p_{next} is about $\log p$, clearly

$$\liminf_{p\to\infty} \frac{p_{\text{next}} - p}{\log p} \le 1.$$

Erdős [9] was the first to show that the lim inf is strictly less than 1. Other landmark results in the area are the works of Bombieri and Davenport [3], Huxley [20], and Maier [22], who introduced several new ideas to this study and progressively reduced the lim inf to $\le 0.24\ldots$. Enter Goldston, Pintz, and Yıldırım:

Theorem 1. *We have*

$$\liminf_{p\to\infty} \frac{p_{\text{next}} - p}{\log p} = 0.$$

So there are substantially smaller gaps between primes than the average! What about even smaller gaps? Can we show that $\liminf_{p\to\infty}(p_{\text{next}} - p) < \infty$ (bounded gaps), or perhaps even $\liminf_{p\to\infty}(p_{\text{next}} - p) = 2$ (twin primes!)?

Theorem 2. *Suppose the Elliott-Halberstam conjecture on the distribution of primes in arithmetic progressions holds true. Then*

$$\liminf_{p\to\infty}(p_{\text{next}} - p) \le 16.$$

What is the Elliott-Halberstam conjecture? One valuable thing that we know about primes is their distribution in arithmetic progressions. Knowledge of this, in the form of the Bombieri-Vinogradov theorem, plays a crucial role in the proof of Theorem 1. To obtain the stronger conclusion of Theorem 2, one needs a better understanding of the distribution of primes in progressions and the Elliott-Halberstam conjecture provides the necessary stronger input. Vaguely, the Goldston-Pintz-Yıldırım results say that if the primes are well separated with no small gaps between them, then something weird must happen to their distribution in progressions.

Given a progression $a \pmod q$ let $\pi(x; q, a)$ denote the number of primes below x lying in this progression. Naturally we may suppose that a and q are coprime, else there is at most one prime in the progression. Now there are $\phi(q)$ — this is Euler's ϕ-function — such progressions $a \pmod q$ with a coprime to q. We would expect that each progression captures its fair share of primes. In other words we expect that $\pi(x; q, a)$ is roughly $\pi(x)/\phi(q)$. The prime number theorem in arithmetic progressions tells us that this is true if we view q as being fixed and let x go to infinity.

In applications, such as Theorem 1, we need information on $\pi(x; q, a)$ when q is not fixed, but growing with x. When q is growing slowly, say q is like $\log x$, the prime number theorem in arithmetic progressions still applies. However if q is a little larger, say q is of size $x^{\frac{1}{3}}$, then currently we cannot prove the equidistribution of primes in the available residue classes $\pmod q$. Such a result would be implied by the Generalized Riemann Hypothesis (indeed for q up to about $\sqrt{x}$), but of course the Generalized Riemann Hypothesis remains unresolved. In this context, Bombieri and Vinogradov showed that the equidistribution of primes in progressions holds, not for each individual q, but on average over q (that is, for a typical q) for q going up to about $\sqrt{x}$. Their result may be thought of as the 'Generalized Riemann Hypothesis on average.'

The Elliott-Halberstam conjecture says that the equidistribution of primes in progressions continues to hold on average for q going up to $x^{1-\epsilon}$ for any given positive ϵ. In some ways, this lies deeper than the Generalized Riemann Hypothesis which permits only $q \le \sqrt{x}$.

We hope that the reader has formed a rough impression of the nature of the assumption in Theorem 2. We will state the Bombieri-Vinogradov theorem and Elliott-Halberstam conjecture precisely in the penultimate section devoted to primes in progressions.

The Hardy-Littlewood conjectures. We already noticed a faulty feature of the Cramér model: given a large prime p, the probability that $p+1$ is prime is not $1/\log(p+1)$ but 0 because $p+1$ is even. Neither would we expect the conditional probability of $p+2$ being prime to be simply $1/\log(p+2)$: after all, $p+2$ is guaranteed to be odd and this should give it a better chance of being prime. How should we formulate the correct probability for $p+2$ being prime? More precisely, what should be the conjectural asymptotics for

$$\#\{p \le x : \ p+2 \text{ prime}\}?$$

The Cramér model would have predicted that this is about $x/(\log x)^2$. While we must definitely modify this, it also seems reasonable that $x/(\log x)^2$ is the right size for the answer. So maybe the answer is about $cx/(\log x)^2$ for an appropriate constant c.

Long ago Hardy and Littlewood [16] figured out what the right conjecture should be. The problem with the Cramér model is that it treats n and $n+2$ as being independent, whereas they are clearly dependent. If we want n and $n+2$ both to be prime, then they must both be odd, neither of them must be divisible by 3, nor by 5, and so on. If we choose n randomly, the probability that n and $n+2$ are both odd is $1/2$. In contrast, two randomly chosen numbers would both be odd with a $1/4$ probability. If neither n nor $n+2$ is divisible by 3 then n must be 2 (mod 3), which has a $1/3$ probability. On the other hand, the probability that two randomly chosen numbers are not divisible by 3 is $(2/3)\cdot(2/3) = 4/9$. Similarly, for any prime $\ell \geq 3$, the probability that n and $n+2$ are not divisible by ℓ is $1-2/\ell$, which is a little different from the probability $(1-1/\ell)^2$ that two randomly chosen integers are both not divisible by ℓ. For the prime 2 we must correct the probability $1/4$ by multiplying by $2 = (1-1/2)(1-1/2)^{-2}$, and for all primes $\ell \geq 3$ we must correct the probability $(1-1/\ell)^2$ by multiplying by $(1-2/\ell)(1-1/\ell)^{-2}$. The idea is that if we multiply all these correction factors together then we have accounted for 'all the ways' in which n and $n+2$ are dependent, producing the required correction constant c. Thus the conjectured value for c is the product over primes

$$\Big(1-\frac{1}{2}\Big)\Big(1-\frac{1}{2}\Big)^{-2}\prod_{\ell\geq 3}\Big(1-\frac{2}{\ell}\Big)\Big(1-\frac{1}{\ell}\Big)^{-2}.$$

Let us make a synthesis of the argument above, which will allow us to generalize it. For any prime ℓ let $\nu_{\{0,2\}}(\ell)$ denote the number of distinct residue classes (mod ℓ) occupied by the numbers 0 and 2. If we want n and $n+2$ to be both coprime to ℓ then n must n must avoid the residue classes occupied by -0 and -2 (mod ℓ), so that n must lie in one of $\ell - \nu_{\{0,2\}}(\ell)$ residue classes. The probability that this happens is $1-\nu_{\{0,2\}}(\ell)/\ell$, so the correction factor for ℓ is $(1-\nu_{\{0,2\}}(\ell)/\ell)(1-1/\ell)^{-2}$. As before, consider the infinite product over primes

$$\mathfrak{S}(\{0,2\}) := \prod_{\ell}\Big(1-\frac{\nu_{\{0,2\}}(\ell)}{\ell}\Big)\Big(1-\frac{1}{\ell}\Big)^{-2}.$$

The infinite product certainly converges: the terms for $\ell \geq 3$ are all less than 1 in size. Moreover, it converges to a non-zero number. Note that none of the factors above is zero, and that for large ℓ the logarithm of the corresponding factor above is very small — it is $\log(1-1/(\ell-1)^2) \approx -1/\ell^2$. Thus the sum of the logarithms converges, and the product is non-zero; indeed $\mathfrak{S}(\{0,2\})$ is numerically about 1.3203. Then the conjecture is that for large x

$$\#\{p \leq x : \ p+2 \text{ prime}\} \sim \mathfrak{S}(\{0,2\})\frac{x}{(\log x)^2}.$$

Here and below, the notation $f(x) \sim g(x)$ means that $\lim_{x\to\infty} f(x)/g(x) = 1$.

The conjecture generalizes readily: Suppose we are given a set $\mathcal{H} = \{h_1, h_2, \dots, h_k\}$ of non-negative integers and we want to find the frequency with which $n+h_1, \dots, n+h_k$

are all prime. For a prime number ℓ, we define $\nu_{\mathcal{H}}(\ell)$ to be the number of distinct residue classes (mod ℓ) occupied by $\mathcal{H}$. We define the 'singular series'[5]

$$\mathfrak{S}(\mathcal{H}) = \prod_{\ell} \Big(1 - \frac{\nu_{\mathcal{H}}(\ell)}{\ell}\Big)\Big(1 - \frac{1}{\ell}\Big)^{-k}. \tag{2}$$

If ℓ is larger than all elements of $\mathcal{H}$ then $\nu_{\mathcal{H}}(\ell) = k$, and for such ℓ the terms in the product are less than 1. Thus the product converges. When does it converge to a non-zero number? If $\nu_{\mathcal{H}}(\ell) = \ell$ for some prime ℓ then one of the terms in our product vanishes, and so our product must be zero. Suppose none of the terms is zero. For large ℓ the logarithm of the corresponding factor is

$$\log\Big(1 - \frac{k}{\ell}\Big)\Big(1 - \frac{1}{\ell}\Big)^{-k} \approx -\frac{k(k+1)}{2\ell^2},$$

and so the sum of the logarithms converges, and our product is non-zero. Thus the singular series is zero if and only if $\nu_{\mathcal{H}}(\ell) = \ell$ for some prime ℓ — that is, if and only if the numbers $h_1, \dots, h_k$ occupy *all* the residue classes (mod ℓ) for some prime ℓ. In that case, for any n one of the numbers $n + h_1, \dots, n + h_k$ must be a multiple of ℓ, and so there are only finitely many prime k-tuples $n + h_1, \dots, n + h_k$.

The Hardy-Littlewood conjecture. *Let $\mathcal{H} = \{h_1, \dots, h_k\}$ be a set of positive integers such that $\mathfrak{S}(\mathcal{H}) \neq 0$. Then*

$$\#\{n \le x : \ n + h_1, \dots, n + h_k \text{ prime}\} \sim \mathfrak{S}(\mathcal{H}) \frac{x}{(\log x)^k}.$$

It is easy to see that $\mathfrak{S}(\{0, 2r\}) \neq 0$ for every non-zero even number $2r$. Thus the Hardy-Littlewood conjecture predicts that there are about $\mathfrak{S}(\{0, 2r\})x/(\log x)^2$ prime pairs p and $p+2r$ with p below x. Further, the number of these pairs for which $p+2d$ is prime for some d between 1 and $r-1$ is at most a constant times $x/(\log x)^3$. We deduce that there should be infinitely many primes p for which the gap to the next prime is exactly $2r$. Thus every positive even number should occur infinitely often as a gap between successive primes, but we don't know this for a single even number!

For any k, it is easy to find k-element sets $\mathcal{H}$ with $\mathfrak{S}(\mathcal{H}) \neq 0$. For example, take $\mathcal{H}$ to be any k primes all larger than k. Clearly if $\ell > k$ then $\nu_{\mathcal{H}}(\ell) \le k < \ell$, while if $\ell \le k$ then the residue class 0 (mod ℓ) must be omitted by the elements of $\mathcal{H}$ (they are primes!) and so once again $\nu_{\mathcal{H}}(\ell) < \ell$.

We make one final comment before turning (at last!) to the ideas behind the proofs of Theorems 1 and 2. Conjecture 1 was made on the strength of the Cramér model, but we have just been discussing how to modify the Cramér probabilities for prime k-tuples. A natural question is whether the Hardy-Littlewood conjectures are consistent

[5]The terminology is not entirely whimsical: Hardy and Littlewood originally arrived at their conjecture through a heuristic application of their 'circle method.' In their derivation, $\mathfrak{S}(\mathcal{H})$ did arise as a series rather than as our product.

with Conjecture 1. In a beautiful calculation [11], Gallagher showed that Conjecture 1 can in fact be obtained starting from the Hardy-Littlewood conjectures. The crucial point in his proof is that although $\mathfrak{S}(\mathcal{H})$ is not always 1 (as the Cramér model would have), it is *approximately* 1 on average over all k-element sets $\mathcal{H}$ with the $h_j \le h$. That is, as $h \to \infty$,

$$\sum_{1\le h_1<h_2<\ldots<h_k\le h} \mathfrak{S}(\{h_1,\ldots,h_k\}) \sim \sum_{1\le h_1<h_2<\ldots<h_k\le h} 1. \tag{3}$$

The ideas of Goldston, Pintz and Yıldırım. We will start with the idea behind Theorem 2. Let k be a given positive integer which is at least 2. Let $\mathcal{H} = \{h_1 < \ldots < h_k\}$ be a set with $\mathfrak{S}(\mathcal{H}) \neq 0$. We aspire to the Hardy-Littlewood conjecture which says that there must be infinitely many n such that $n + h_1, \ldots, n + h_k$ are all prime. Since there are infinitely many primes, trivially at least one of the numbers $n + h_1, \ldots, n + h_k$ is prime infinitely often. Can we do a little better: can we show that two of the numbers $n+h_1, \ldots, n+h_k$ are prime infinitely often? If we could, then we would plainly have that $\liminf_{p\to\infty}(p_{\text{next}} - p) \le (h_k - h_1)$.

How do we detect two primes in $n + h_1, \ldots, n + h_k$? Let x be large and consider n varying between x and $2x$. Suppose we are able to find a function $a(n)$ which is always non-negative, and such that, for each $j = 1, \ldots, k$,

$$\sum_{\substack{x\le n\le 2x\\ n+h_j \text{ prime}}} a(n) > \frac{1}{k}\sum_{x\le n\le 2x} a(n). \tag{4}$$

Then summing over $j = 1, \ldots, k$, it would follow that

$$\sum_{x\le n\le 2x} \#\{1 \le j \le k : n + h_j \text{ prime}\}\; a(n) > \sum_{x\le n\le 2x} a(n),$$

so that for some number n lying between x and $2x$ we must have at least two primes among $n + h_1, \ldots, n + h_k$.

Of course, the question is how do we find such a function $a(n)$ satisfying (4)? We would like to take $a(n) = 1$ if $n + h_1, \ldots, n + h_k$ are all prime, and 0 otherwise. But then evaluating the problem of evaluating $\sum_{x\le n\le 2x} a(n)$ is precisely that of establishing the Hardy-Littlewood conjecture.

The answer is suggested by sieve theory, especially the theory of Selberg's sieve. Sieve theory is concerned with finding primes, or numbers without too many prime factors, among various integer sequences. Some of the spectacular achievements of this theory are Chen's theorem [5] that for infinitely many primes p, the number $p + 2$ has at most two prime factors; the result of Friedlander and Iwaniec [10] that there are infinitely many primes of the form $x^2 + y^4$; and the result of Heath-Brown [19] that there are infinitely many primes of the form $x^3 + 2y^3$. We recall here very briefly the idea behind Selberg's sieve.

Interlude on Selberg's sieve. We illustrate Selberg's sieve by giving an upper bound on the number of prime k-tuples $n+h_1, \ldots, n+h_k$ with $x \le n \le 2x$. The idea is to find a

'nice' function $a(n)$ which equals 1 if $n+h_1, \dots, n+h_k$ are all prime, and is non-negative otherwise. Then $\sum_{x\le n\le 2x} a(n)$ provides an upper bound for the number of prime k-tuples. Of course, we must choose $a(n)$ appropriately, so as to be able to evaluate $\sum_{x\le n\le 2x} a(n)$.

Selberg's choice for $a(n)$ is as follows: Let λ_d be a sequence of real numbers such that

$$\lambda_1 = 1, \qquad \text{and with} \qquad \lambda_d = 0 \qquad \text{for } d > R. \tag{5}$$

Choose[6]

$$a(n) = \Big(\sum_{d \mid (n+h_1)\dots(n+h_k)} \lambda_d \Big)^2. \tag{6}$$

Being a square, $a(n)$ is clearly non-negative. If $R < x \le n$ and $n+h_1, \dots, n+h_k$ are all prime, then the only non-zero term in (6) is for $d = 1$ and so $a(n) = 1$ as desired. Therefore we assume that $R < x$ below. The goal is to choose λ_d so as to minimize $\sum_{x\le n\le 2x} a(n)$. There is an advantage to allowing R as large as possible, since this gives us greater flexibility in choosing the parameters λ_d. On the other hand it is easier to estimate $\sum_{x\le n\le 2x} a(n)$ when R is small since there are fewer divisors d to consider. In the problem at hand, it turns out that we can choose R roughly of size $\sqrt{x}$. This choice leads to an upper bound for the number of prime k-tuples of about $2^k \cdot k! \mathfrak{S}(\mathcal{H}) x/(\log x)^k$. That is, a bound about $2^k \cdot k!$ times the conjectured Hardy-Littlewood asymptotic.

Expanding out the square in (6) and summing over n, we must evaluate

$$\sum_{d_1,d_2} \lambda_{d_1}\lambda_{d_2} \sum_{\substack{x\le n\le 2x \\ d_1 \mid (n+h_1)\cdots(n+h_k) \\ d_2 \mid (n+h_1)\cdots(n+h_k)}} 1 = \sum_{d_1,d_2} \lambda_{d_1}\lambda_{d_2} \sum_{\substack{x\le n\le 2x \\ [d_1,d_2] \mid (n+h_1)\cdots(n+h_k)}} 1,$$

where $[d_1, d_2]$ denotes the l.c.m. of d_1 and d_2. The condition $[d_1, d_2] | (n+h_1)\cdots(n+h_k)$ means that n must lie in a certain number (say, $f([d_1, d_2])$) of residue classes $\pmod{[d_1, d_2]}$. Can we count the number of $x \le n \le 2x$ lying in the union of these arithmetic progressions? Divide the interval $[x, 2x]$ into intervals of length $[d_1, d_2]$ with possibly one smaller interval left over at the end. Each complete interval (and there are about $x/[d_1, d_2]$ of these) gives $f([d_1, d_2])$ values of n; the last shorter interval contributes an indeterminate 'error' between 0 and $f([d_1, d_2])$. So, at least if $[d_1, d_2]$ is a bit smaller than x, we can estimate the sum over n accurately. Since $[d_1, d_2] \le d_1 d_2 \le R^2$, if R is a bit smaller than[7] $\sqrt{x}$, then the sum over n can be evaluated accurately. Let us suppose that R is about size $\sqrt{x}$ and that the error terms can be disposed of satisfactorily. It remains to handle the main term contribution to $\sum_{x\le n\le 2x} a(n)$, namely

$$x \sum_{d_1,d_2\le R} \frac{f([d_1, d_2])}{[d_1, d_2]} \lambda_{d_1}\lambda_{d_2}. \tag{7}$$

[6]Below, the symbol $a|b$ means that a divides b.

[7]To be precise, R must be $\le \sqrt{x}/(\log x)^{2k}$, say.

The reader may wonder what $f([d_1, d_2])$ is. Let us work this out in the case when $[d_1, d_2]$ is not divisible by the square of any prime; the other case is more complicated, but not very important in this problem. If p is a prime and we want $p|(n+h_1)\cdots(n+h_k)$ then clearly $n \equiv -h_j \pmod p$ for some j, so that n lies in one of $\nu_{\mathcal{H}}(p)$ residue classes $\pmod p$. By the chinese remainder theorem it follows that if $[d_1, d_2]|(n+h_1)\cdots(n+h_k)$ then n lies in $\prod_{p|[d_1,d_2]} \nu_{\mathcal{H}}(p)$ residue classes $\pmod{[d_1, d_2]}$. Thus f is a multiplicative function[8], with $f(p) = \nu_{\mathcal{H}}(p)$.

The problem in Selberg's sieve is to choose λ_d subject to the linear constraint (5) in such a way as to minimize the quadratic form (7) (that would give the best upper bound for $\sum_{x\le n\le 2x} a(n)$). This can be achieved using Lagrange multipliers, or by diagonalizing the quadratic form (7). We do not give the details of this calculation but just record the result obtained. The optimal choice of λ_d for $d \le R$ is given by

$$\lambda_d \approx \mu(d)\Big(\frac{\log R/d}{\log R}\Big)^k,$$

where $\mu(d)$ is the Möbius function.[9] With this choice of λ_d the quantity in (7) is

$$\approx k!\mathfrak{S}(\mathcal{H})\frac{x}{(\log R)^k} \approx 2^k \cdot k!\mathfrak{S}(\mathcal{H})\frac{x}{(\log x)^k}.$$

The appearance at this stage of the Möbius function is not surprising, as it is very intimately connected with primes. For example, the reader can check that $\sum_{d|m} \mu(d)(\log m/d)^k$ equals 0 unless m is divisible by at most k distinct prime factors. When $m = p_1\cdots p_k$ is the product of k distinct prime factors it equals $k!(\log p_1)\cdots(\log p_k)$, and there is a more complicated formula if m is composed of fewer than k primes, or if m is divisible by powers of primes. Applying this to $m = (n+h_1)\cdots(n+h_k)$, we are essentially picking out prime k-tuples! The optimum in Selberg's sieve is a kind of approximation to this identity.

Return to Goldston-Pintz-Yıldırım. We want to find a non-negative function $a(n)$ so as to make (4) hold. Motivated by Selberg's sieve we may try to find optimal λ_d as in (5) and again choose $a(n)$ as in (6). If we try such a choice, then our problem now is to maximize the ratio

$$\Big(\sum_{\substack{x\le n\le 2x\\ n+h_j \text{prime}}} a(n)\Big)\Big/\Big(\sum_{x\le n\le 2x} a(n)\Big). \tag{8}$$

We'd like this ratio to be $> 1/k$. Notice again that it is advantageous to choose R as large as possible to give greatest freedom in choosing λ_d, but in order to evaluate the sums above there may be restrictions on the size of R. In dealing with the denominator we saw that there is a restriction $R \le \sqrt{x}$ (essentially) and that in this situation the denominator in (8) is given by the quadratic form (7). We will see below that in dealing with the numerator

[8]These are functions satisfying $f(mn) = f(m)f(n)$ for any pair of coprime integers m and n.

[9]$\mu(d) = 0$ if d is divisible by the square of a prime. Otherwise $\mu(d) = (-1)^{\omega(d)}$ where $\omega(d)$ is the number of distinct primes dividing d.

of (8), a more stringent restriction on R must be made: we can only take R around size $x^{\frac{1}{4}}$.

In any case, (8) is the ratio of two quadratic forms, and this ratio needs to be maximized keeping in mind the linear constraint (5). This optimization problem is more delicate than the one in Selberg's sieve. It is not clear how to proceed most generally: Lagrange multipliers become quite messy, and we can't quite diagonalize both quadratic forms simultaneously. It helps to narrow the search to a special class of λ_d. Motivated by Selberg's sieve we will search for the optimum among the choices (for $d \le R$)

$$\lambda_d = \mu(d) P\Big(\frac{\log R/d}{\log R}\Big).$$

Here $P(y)$ denotes a polynomial such that $P(1) = 1$ and such that P vanishes to order at least k at $y = 0$. The condition that P be a polynomial can be relaxed a bit but this is not important. It is however vital for the analysis that P should vanish to order k at 0. Our aim is to find a choice for P which makes the ratio in (8) large.

With this choice of λ_d we can use standard arguments to evaluate (7) and thus the denominator in (8). Omitting the long, technical details, the answer is that for R a little below $\sqrt{x}$, the denominator in (8) is

$$\sim \frac{x}{(\log R)^k} \mathfrak{S}(\mathcal{H}) \int_0^1 \frac{y^{k-1}}{(k-1)!} P^{(k)}(1-y)^2 dy, \tag{9}$$

where $P^{(k)}$ denotes the k-th derivative of the polynomial P.

To handle the numerator of (8), we expand out the square in (6) and sum over $x \le n \le 2x$ with $n + h_j$ being prime. Thus the numerator is

$$\sum_{d_1, d_2 \le R} \lambda_{d_1} \lambda_{d_2} \sum_{\substack{x \le n \le 2x \\ [d_1,d_2] | (n+h_1)\cdots(n+h_k) \\ n+h_j \text{ prime}}} 1.$$

How can we evaluate the inner sum over n? As we saw before, the condition $[d_1, d_2]$ divides $(n+h_1)\cdots(n+h_k)$ means that n lies in $f([d_1, d_2])$ arithmetic progressions $\pmod{[d_1, d_2]}$. For each of these progressions we must count the number of n such that $n + h_j$ is prime. Of course, for some of the $f([d_1, d_2])$ progressions it may happen that $n + h_j$ automatically has a common factor with $[d_1, d_2]$ and so cannot be prime. Suppose there are $g([d_1, d_2])$ progressions such that $n + h_j$ is guaranteed to be coprime to $[d_1, d_2]$. For each of these progressions we are counting the number of primes between x and $2x$ lying in a reduced residue class[10] $\pmod{[d_1, d_2]}$. Given a modulus q, the prime number theorem in arithmetic progressions says that the primes are roughly equally divided among the reduced residue classes $\pmod q$. Thus, ignoring error terms completely, we expect the sum over n to be about

$$\frac{\pi(2x) - \pi(x)}{\phi([d_1, d_2])} g([d_1, d_2]).$$

[10]A reduced residue class $\pmod q$ is a progression $a \pmod q$ where a is coprime to q.

The $\phi([d_1, d_2])$ in the denominator is Euler's ϕ-function: for any integer m, $\phi(m)$ counts the number of reduced residue classes $\pmod m$. Since $\pi(2x) - \pi(x)$ is about $x/\log x$ we 'conclude' that the numerator in (8) is about

$$\frac{x}{\log x} \sum_{d_1, d_2 \leq R} \lambda_{d_1} \lambda_{d_2} \frac{g([d_1, d_2])}{\phi([d_1, d_2])}. \tag{10}$$

This is the expression analogous to (7) for the numerator.

Two big questions: what is the function g, and for what range of R can we handle the error terms above? Let us first describe g. As with f let us suppose that $[d_1, d_2]$ is not divisible by the square of any prime. As noted earlier, if p is prime and $p|(n+h_1)\cdots(n+h_k)$ then n lies in one of $\nu_{\mathcal{H}}(p)$ residue classes $\pmod p$. If we want $n + h_k$ to be prime, then one of these residue classes, namely $n \equiv -h_j \pmod p$, must be forbidden. Thus there are now $\nu_{\mathcal{H}}(p) - 1$ residue classes available for $n \pmod p$. In other words, $g(p) = \nu_{\mathcal{H}}(p) - 1$, and the chinese remainder theorem shows that g must be defined multiplicatively:

$$g([d_1, d_2]) = \prod_{p|[d_1,d_2]} (\nu_{\mathcal{H}}(p) - 1).$$

We will postpone the detailed discussion on primes in arithmetic progressions which is needed to handle the error terms above. For the moment, let us note that the Bombieri-Vinogradov theorem (which is a powerful substitute for the generalized Riemann hypothesis in many applications) allows us to control $\pi(x; q, a)$ (the number of primes up to x which are congruent to $a \pmod q$), on average over q, for q up to about $\sqrt{x}$. Since our moduli are $[d_1, d_2]$, which go up to R^2, we see that R may be chosen up to about $x^{\frac{1}{4}}$. Conjectures of Montgomery, and Elliott and Halberstam (discussed below) would permit larger values of R, going up to $x^{\frac{1}{2}-\epsilon}$ for any $\epsilon > 0$.

Thus, with R a little below $x^{\frac{1}{4}}$, the expression (10) does give a good approximation to the numerator of (8). Now a standard but technical argument can be used to evaluate (10). As with (9), the answer is

$$\sim \frac{x}{(\log x)(\log R)^{k-1}} \mathfrak{S}(\mathcal{H}) \int_0^1 \frac{y^{k-2}}{(k-2)!} P^{(k-1)}(1-y)^2 dy. \tag{11}$$

Assuming that $\mathfrak{S}(\mathcal{H}) \neq 0$, it follows from (9) and (11) that the ratio in (8) is about

$$\frac{\log R}{\log x} \Big(\int_0^1 \frac{y^{k-2}}{(k-2)!} P^{(k-1)}(1-y)^2 dy \Big) \Big/ \Big(\int_0^1 \frac{y^{k-1}}{(k-1)!} P^{(k)}(1-y)^2 dy \Big). \tag{12}$$

This is the moment of truth: can we choose P so as to make this a little larger than $1/k$?

Here is a good choice for P: take $P(y) = y^{k+r}$ for a non-negative integer r to be chosen optimally. After some calculations with beta-integrals, we see that (12) then equals

$$\Big(\frac{\log R}{\log x}\Big)\Big(\frac{2(2r+1)}{(r+1)(k+2r+1)}\Big).$$

This is largest when r is about $\sqrt{k}/2$, and the second fraction above is close to but less than $4/k$. Since we can choose R a little below $x^{\frac14}$, the first fraction is close to but less than $1/4$. Thus (12) is very close to, but less than, $1/k$. We therefore barely fail to prove bounded gaps between primes! Of course, we just tried one choice of P; maybe there is a better choice which gets us over the edge. Unfortunately, the second fraction in (12) *cannot* be made larger than $4/k$. If we set $Q(y) = P^{(k-1)}(y)$ then Q is a polynomial, not identically zero, with $Q(0) = 0$; for such polynomials Q we claim that the unfortunate inequality

$$\int_0^1 \frac{y^{k-2}}{(k-2)!}Q(1-y)^2dy < \frac{4}{k}\int_0^1 \frac{y^{k-1}}{(k-1)!}Q'(1-y)^2dy$$

holds. The reader can try her hand at proving this.

We now have enough to prove Theorem 2! If we can choose R a little larger than $x^{\frac14}$ then for suitably large k the quantity in (12) can be made larger than $1/k$ as desired. If we allow $R = x^{\frac12-\epsilon}$ as the Elliott-Halberstam conjecture predicts, then with $k = 7$ and $r = 1$ we can make (12) nearly $1.05/k > 1/k$. Thus, if we take any set $\mathcal{H}$ with seven elements and $\mathfrak{S}(\mathcal{H}) \neq 0$ then for infinitely many n at least two of the numbers $n + h_1, \dots, n + h_k$ are prime! By choosing a more careful polynomial P we can make do with six element sets $\mathcal{H}$ rather than seven. The first six primes larger than 6 are $7, 11, 13, 17, 19$, and 23, and so $\mathfrak{S}(\{7, 11, 13, 17, 19, 23\}) \neq 0$. Thus, it follows that — assuming the Elliott-Halberstam conjecture — there are infinitely many gaps between primes that are at most 16.

What can we recover unconditionally? We are so close to proving Theorem 2 unconditionally, that clearly some tweaking of the argument must give Theorem 1! The idea here is to average over sets $\mathcal{H}$. For clarity, let us now denote $a(n)$ above by $a(n; \mathcal{H})$ to exhibit the dependence on $\mathcal{H}$.

Given $\epsilon > 0$ we wish to find primes p between x and $2x$ such that $p_{\text{next}} - p \le \epsilon \log x$. This would prove Theorem 1. Set $h = \epsilon \log x$, and let k be a natural number chosen in terms of ϵ, but fixed compared to x. Consider the following two sums:

$$\sum_{1\le h_1<h_2<\dots<h_k\le h} \sum_{x\le n\le 2x} a(n; \{h_1, \dots, h_k\}), \tag{13}$$

and

$$\sum_{1\le h_1<h_2<\dots<h_k\le h} \sum_{1\le \ell\le h} \sum_{\substack{x\le n\le 2x\\ n+\ell \text{ prime}}} a(n; \{h_1, \dots, h_k\}). \tag{14}$$

If we could prove that (14) is larger than (13), it would follow that for some n between x and $2x$, there are two prime numbers between $n + 1$ and $n + h$, as desired.

Our analysis above already gives us the asymptotics for (13) and (14). Using (9) we see that the quantity (13) is

$$\sim \frac{x}{(\log R)^k}\Big(\int_0^1 \frac{y^{k-1}}{(k-1)!}P^{(k)}(1-y)^2dy\Big) \sum_{1\le h_1<h_2<\dots<h_k\le h} \mathfrak{S}(\{h_1, \dots, h_k\}),$$

and using Gallagher's result (3) this is

$$\sim \frac{x}{(\log R)^k}\frac{h^k}{k!}\int_0^1 \frac{y^{k-1}}{(k-1)!}P^{(k)}(1-y)^2dy. \tag{15}$$

Now let us consider (14). Here we distinguish two cases: the case when $\ell = h_j$ for some j, and the case when $\ell \neq h_j$ for all j. The former case is handled by our analysis leading up to (11). Upon using (3) again, these terms contribute

$$\sim k\frac{x}{(\log x)(\log R)^{k-1}}\frac{h^k}{k!}\int_0^1 \frac{y^{k-2}}{(k-2)!}P^{(k-1)}(1-y)^2dy. \tag{16}$$

If we choose $P(y) = y^{k+r}$ as before, we see that (16) is already just a shade below (15), so we need the slightest bit of extra help from the terms $\ell \neq h_j$ for any j. If $n+\ell$ is prime note that

$$\begin{aligned} a(n;\{h_1,\dots,h_k\}) &= \Big(\sum_{d|(n+h_1)\cdots(n+h_k)}\lambda_d\Big)^2 \\ &= \Big(\sum_{d|(n+h_1)\cdots(n+h_k)(n+\ell)}\lambda_d\Big)^2 = a(n;\{h_1,\dots,h_k,\ell\}), \end{aligned}$$

since the divisors counted in the latter sum but not the former are all larger than $n+\ell > x > R$ and so $\lambda_d = 0$ for such divisors. This allows us to finesse the calculation by simply appealing to (11) again, with k replaced by $k+1$ and $\{h_1,\dots,h_k\}$ by $\{h_1,\dots,h_k,\ell\}$. Thus the latter class of integers ℓ contributes

$$\sim \sum_{1\le h_1<h_2<\ldots<h_k\le h}\sum_{\substack{\ell=1\\ \ell\neq h_j}}^{h}\frac{x}{(\log x)(\log R)^k}\mathfrak{S}(\{h_1,\dots,h_k,\ell\})\int_0^1\frac{y^{k-1}}{(k-1)!}P^{(k)}(1-y)^2dy.$$

Appealing to (3) again — we are now summing over $k+1$ element sets but each set is counted $k+1$ times — this is

$$\sim \frac{x}{(\log R)^k}\frac{h^k}{k!}\frac{h}{\log x}\int_0^1\frac{y^{k-1}}{(k-1)!}P^{(k)}(1-y)^2dy. \tag{17}$$

This accounts for a factor of ϵ times the quantity in (15), and now the combined contribution of (16) and (17) may be made larger than (15), proving Theorem 1!

Primes in arithmetic progressions. It remains to explain what is meant by the Bombieri-Vinogradov theorem and the Elliott-Halberstam conjecture. Recall that we required knowledge of these estimates for primes in progressions while discussing the error terms that arise while evaluating the numerator of (8).

Let us write

$$\pi(x) = \mathrm{li}(x) + E(x),$$

where $\mathrm{li}(x)$ stands for the 'logarithmic integral' $\int_2^x \frac{dt}{\log t}$, which is the expected main term, and $E(x)$ stands for an 'error term'. The main term $\mathrm{li}(x)$ is, by integration by parts, roughly $x/\log x$. As for the error term $E(x)$, the standard proofs of the Prime Number Theorem give that for any number $A > 0$ there exists a constant $C(A)$ such that

$$|E(x)| \le C(A)\frac{x}{(\log x)^A}.$$

The argument generalizes readily for primes in progressions. Given an arithmetic progression $a \pmod q$ with $(a,q) = 1$ let us write

$$\pi(x;q,a) = \frac{1}{\phi(q)}\mathrm{li}(x) + E(x;q,a),$$

where $\mathrm{li}(x)/\phi(q)$ is the expected main term — the primes are equally divided among the available residue classes — and $E(x;q,a)$ is an 'error term' which we would like to be small. As with the Prime Number Theorem, for every $A > 0$ there exists a constant $C(q,A)$ such that

$$|E(x;q,a)| \le C(q,A)\frac{x}{(\log x)^A}.$$

We emphasize that the constant $C(q,A)$ may depend on q. Therefore, this result is meaningful only if we think of q as being fixed and let x tend to ∞. In applications such a result is not very useful, because we may require q not to be fixed, but to grow with x. For example, in our discussions above we want to deal with primes in progressions $\pmod{[d_1,d_2]}$ which can be as large as R^2, and we'd like this to be of size $x^{\frac12}$ and would love to have it be even larger. Thus the key issue while discussing primes in arithmetic progressions is the uniformity in q with which the asymptotic formula holds.

What is known about $\pi(x;q,a)$ for an individual modulus q is disturbingly weak. From a result of Siegel we know that for any given positive numbers N and A, there exists a constant $c(N,A)$ such that if $q < (\log x)^N$ then

$$|E(x;q,a)| \le c(N,A)\frac{x}{(\log x)^A}.$$

This is better than the result for fixed q mentioned earlier, but the range of q is still very restrictive. An additional defect is that the constant $c(N,A)$ cannot be computed explicitly[11] in terms of N and A.

If we assume the Generalized Riemann Hypothesis (GRH) then we would fare much better: if $x \ge q$ there exists a positive constant C independent of q such that

$$|E(x;q,a)| \le Cx^{\frac12}\log x.$$

This gives a good asymptotic formula for $\pi(x;q,a)$ in the range $q \le x^{\frac12}/(\log x)^3$, say.

[11]This is not due to laziness, but is a fundamental defect of the method of proof.

Given a modulus q let us define

$$E(x;q) = \max_{(a,q)=1} |E(x;q,a)|.$$

We have discussed above the available weak bounds for $E(x;q)$, and the unavailable strong GRH bound. Luckily, in many applications including ours, we don't need a bound for $E(x;q)$ for each individual q, but only a bound holding in an average sense as q varies. In the application to small gaps, we want primes in progressions $\pmod{[d_1,d_2]}$, but recall that we also have a sum over d_1, d_2 going up to R. An extremely powerful result of Bombieri and Vinogradov gives such an average estimate for $E(x;q)$. Moreover, this average result is nearly as good as what would be implied by the GRH.

The Bombieri-Vinogradov theorem. *For any positive constant A there exist constants B and C such that*

$$\sum_{q\le\mathcal{Q}} \max_{y\le x} |E(y;q)| \le C\frac{x}{(\log x)^A}, \tag{18}$$

with $\mathcal{Q} = x^{\frac12}/(\log x)^B$.

The constant B can be computed explicitly; for example $B = 24A + 46$ is permissible, but the constant C here cannot be computed explicitly (a defect arising from Siegel's theorem mentioned above). The Bombieri-Vinogradov theorem tells us that on average over $q \le \mathcal{Q}$ we have $E(x;q) \le Cx(\log x)^{-A}/\mathcal{Q} = Cx^{\frac12}(\log x)^{B-A}$. Apart from the power of $\log x$, this is as good as the GRH bound!

A straight-forward application of the Bombieri-Vinogradov theorem shows that as long as $R^2 \le x^{\frac12}/(\log x)^B$ for suitably large B, the error terms arising in the Goldston-Pintz-Yıldırım argument will be manageable. If we wish to take R larger, then we must extend the range of $\mathcal{Q}$ in (18). Such extensions are conjectured to hold, but unconditionally the range in (18) has never been improved upon[12].

The Elliott-Halberstam conjecture. *Given $\epsilon > 0$ and $A > 0$ there exists a constant C such that*

$$\sum_{q\le\mathcal{Q}} \max_{y\le x} |E(x;q)| \le C\frac{x}{(\log x)^A},$$

with $\mathcal{Q} = x^{1-\epsilon}$.

The Elliott-Halberstam conjecture would allow us to take $R = x^{\frac12-\epsilon}$ in the Goldston-Pintz-Yıldırım argument. It is worth emphasizing that knowing (18) for $\mathcal{Q} = x^\theta$ with any $\theta > \frac12$ would lead to the existence of bounded gaps between large primes.

Finally, let us mention a conjecture of Montgomery which lies deeper than the GRH and also implies the Elliott-Halberstam conjecture.

[12]Although, Bombieri, Friedlander and Iwaniec [4] have made important progress in related problems

Montgomery's conjecture. *For any $\epsilon > 0$ there exists a constant $C(\epsilon)$ such that for all $q \le x$ we have*

$$E(x;q) \le C(\epsilon) x^{\frac{1}{2}+\epsilon} q^{-\frac{1}{2}}.$$

We have given a very rapid account of prime number theory. For more detailed accounts we refer the reader to the books of Bombieri [2], Davenport [6], and Montgomery and Vaughan [24].

Future directions. We conclude the article by mentioning a few questions related to the work of Goldston-Pintz-Yıldırım.

First and most importantly, is it possible to prove unconditionally the existence of bounded gaps between primes? As it stands, the answer appears to be no, but perhaps suitable variants of the method will succeed. There are other sieve methods available beside Selberg's. Does modifying one of these (e.g. the combinatorial sieve) lead to a better result? If instead of primes we consider numbers with exactly two prime factors, then Goldston, Graham, Pintz, and Yıldırım [13] have shown that there are infinitely many bounded gaps between such numbers.

In a related vein, assuming the Elliott-Halberstam conjecture, can one get to twin primes? Recall that under that assumption, we could show that infinitely many permissible 6-tuples contain two primes. Can the 6 here be reduced? Hopefully, to 2? Again the method in its present form cannot be pushed to yield twin primes, but maybe only one or two new ideas are needed.

Given any $\epsilon > 0$, Theorem 1 shows that for infinitely many n the interval $[n, n + \epsilon \log n]$ contains at least two primes. Can we show that such intervals sometimes contain three primes? Assuming the Elliott-Halberstam conjecture one can get three primes in such intervals, see [12]. Can this be made unconditional? What about k primes in such intervals for larger k?

Is there a version of this method which can be adapted to give long gaps between primes? That is, can one attack Erdős's $10,000 question?

Acknowledgments. I am very grateful to Carine Apparicio, Bryden Cais, Brian Conrad, Sergey Fomin, Andrew Granville, Leo Goldmakher, Rizwan Khan, Jeff Lagarias, Youness Lamzouri, János Pintz, and Trevor Wooley for their careful reading of this article, and many valuable comments.

References

[1] E. Bogomolny and J. Keating, *Random matrix theory and the Riemann zeros. II. n-point correlations*, Nonlinearity **9** (1996), 911–935.
[2] E. Bombieri, *Le grand crible dans la théorie analytique des nombres*, vol. 18, Astérisque, 1987/1974.
[3] E. Bombieri and H. Davenport, *Small differences between prime numbers*, Proc. Roy. Soc. Ser. A **293** (1966), 1–18.
[4] E. Bombieri, J. Friedlander and H. Iwaniec, *Primes in arithmetic progressions to large moduli*, Acta Math. **156** (1986), 203–251.
[5] J.R. Chen, *On the representation of a large even number as the sum of a prime and the product of at most two primes*, Kexue Tongbao (Foreign Lang. Ed.) **17** (1966), 385–386.
[6] H. Davenport, *Multiplicative number theory*, Springer Verlag, 2000.

[7] N. Elkies and C. McMullen, *Gaps in $\sqrt{n}$ (mod 1) and ergodic theory*, Duke Math. J. **123** (2004), 95–139.
[8] P. Erdős, *On the difference of consecutive primes*, Quart. J. Math. Oxford **6** (1935), 124–128.
[9] P. Erdős, *The difference between consecutive primes*, Duke Math. J. **6** (1940), 438–441.
[10] J. Friedlander and H. Iwaniec, *The polynomial X^2+Y^4 captures its primes*, Ann. of Math. **148** (1998), 945–1040.
[11] P. X. Gallagher, *On the distribution of primes in short intervals*, Mathematika **23** (1976), 4–9.
[12] D. Goldston, J. Pintz and C. Yıldırım, *Primes in tuples, I*, preprint, available at www.arxiv.org.
[13] D. Goldston, S. Graham, J. Pintz and C. Yıldırım, *Small gaps between primes and almost primes*, preprint, available at www.arxiv.org.
[14] D. Goldston, Y. Motohashi, J. Pintz and C. Yıldırım, *Small gaps between primes exist*, preprint, available at www.arxiv.org.
[15] A. Granville, *Unexpected irregularities in the distribution of prime numbers*, Proc. of the Int. Congr. of Math., Vol. 1, 2 (Zürich, 1994) (1995), Birkhäuser, Basel, 388-399.
[16] G.H. Hardy and J.E. Littlewood, *Some problems of Partitio Numerorum (III): On the expression of a number as a sum of primes*, Acta Math. **44** (1922), 1–70.
[17] D.R. Heath-Brown, *Prime twins and Siegel zeros*, Proc. London Math. Soc. **47** (1983), 193–224.
[18] D.R. Heath-Brown, *Differences between consecutive primes*, Jahresber. Deutsch. Math.-Verein. **90** (1988), 71–89.
[19] D.R. Heath-Brown, *Primes represented by $x^3 + 2y^3$*, Acta Math. **186** (2001), 1–84.
[20] M. Huxley, *Small differences between consecutive primes. II.*, Mathematika **24** (1977), 142–152.
[21] N. Katz and P. Sarnak, *Zeros of zeta functions and symmetry*, Bull. Amer. Math. Soc. **36** (1999), 1–26.
[22] H. Maier, *Small differences between prime numbers*, Michigan Math. J. **35** (1988), 323–344.
[23] H. Montgomery, *The pair correlation of zeros of the zeta-function*, Proc. Symp. Pure Math. **24** (1972), 181–193.
[24] H. Montgomery and R.C. Vaughan, *Multiplicative number theory I: Classical theory*, Cambridge University Press, 2006.
[25] R. Rankin, *The difference between consecutive primes*, J. London Math. Soc. **13**, 242–244.
[26] E. Szemeredi, *On sets of integers containing no k elements in arithmetic progression*, Proc. International Congress of Math. (Vancouver) **2** (1975), 503–505.
[27] E. Westzynthius, *Über die Verteilung der Zahlen, die zu der n ersten Primzahlen teilerfremd sind*, Comm. Phys. Math. Helsingfors **25** (1931), 1–37.

Department of Mathematics, University of Michigan, Ann Arbor, MI 48109, USA
E-mail address: ksound@umich.edu

Probabilistically Checkable Proofs: A Primer

Madhu Sudan*

December 4, 2005

Abstract

Probabilistically checkable proofs are proofs that can checked probabilistically by reading very few bits of the proof. Roughly ten years back it was shown that proofs could be made probabilistically checkable with a modest increase in their size. While the initial proofs were a little too complex, a recent proof due to Irit Dinur gives a dramatically simple (and radically new) construction of probabilistically checkable proofs. This article explains the notion, presents the formal definition and then introduces the reader to Dinur's work and explains some of the context (but does not reproduce Dinur's proof).

1 Introduction

As advances in mathematics continue at the current rate, editors of mathematical journals increasingly face the challenge of reviewing increasingly long, and often wrong, "proofs" of classical conjectures. Often, even when it is a good guess that a given submission is erroneous, it takes excessive amounts of effort on the editor/reviewer's part to find a specific error one can point to. Most reviewers assume this is an inevitable consequence of the notion of verifying submissions; and expect the complexity of the verification procedure to grow with the length of the submission. The purpose of this article is to point out that this is actually not the case: There does exist a format in which we can ask for proofs of theorems to be written. This format allows for perfectly valid proofs of correct theorems, while any purported proof of an incorrect assertion will be "evidently wrong" (in a manner to be clarified below). We refer to this format of writing proofs as Probabilistically Checkable Proofs (PCPs).

In order to formalize the notion of a probabilistically checkable proof, we start with a bare-bones (computationally simplified) view of logic. A system of logic is described by a collection of axioms which include some "atomic axioms" and some derivation rules. An *assertion* is a sentence, which is simply a sequence of letters over the underlying alphabet. A *proof* of a given assertion is a sequence of sentences ending with the assertion, where each sentence is either one of the axioms or is obtained by applying the derivation rules to the previous sentences in the proof. An assertion which has a proof is a *theorem*. We will use the phrase *argument* to refer to sequence of sentences (which may be offered as "proofs" of "assertions" but whose correctness has not been verified).

*CS & AI Laboratory (CSAIL), Massachusetts Institute of Technology, 32-G640, 32 Vassar Street, Cambridge, MA 02139, USA. http://theory.csail.mit.edu/~madhu. This article supported in part by NSF Award CCR-0312575. Views expressed in this article are those of the author, and not endorsed by NSF.

While systems of logic come in many flavors and allow varying degrees of power in their inference rules and the nature of intermediate sentences that they would allow, the "computational perspective" unifies all of these by using the following abstraction: It suggests that a system of logic is given by a computationally *efficient* algorithm called the *verifier*. The inputs to a verifier is a pair of sequences over some finite alphabet, an assertion T and evidence Π and accepts this pair if and only if Π forms a proof of T in its system of logic. Such verifiers certainly capture all known systems of logic. Indeed without the computational efficiency restriction, it would be impossible to capture the spirit that theorems are often *hard* to prove, but once their proofs are given, they are *easy* to verify. For our purposes, we associate the word "efficient" with the feature that the algorithm runs in time polynomial in the length of its inputs. (As an aside, we note that this distinction between the proving theorems and verifying proofs is currently a conjecture, and is exactly the question examined under the label "Is P=NP?".)

The notion that a verifier can perform any polynomial time computation enriches the class of theorems and proofs considerably and starts to offer highly non-trivial methods of proving theorems. (One immediate consequence is that we can assume theorems/proofs/assertions/arguments are *binary* sequences and we will do so henceforth.) For instance, suppose we have an assertion A (say the Riemann Hypothesis), and say we believe that it has proof which would fit within a 10,000 page article. The computational perspective says that given A and this bound (10,000 pages), one can efficiently compute three positive integers N, L, U with $L \leq U \leq N$ such that A is true if and only if N has a divisor between L and U. The integers N, L, and U will be quite long (maybe writing them would take a million pages), yet they can be produced extremely efficiently (in less than the amount of time it would take a printer to print out all these integers, which is certainly at most a day or two). (This example is based on a result due to Joe Kilian, personal communication.) The theory of NP-completeness could be viewed as an enormous accumulation of many other equivalent formats for writing theorems and proofs. Depending on one's perspective, this may or may not be a better format for writing theorems and proofs. What is important for us is that despite the fact that it differ radically from our mental picture of theorems/proofs - this is as valid a method as any. Every theorem has a valid proof, and this proof in only polynomially larger than the proof in any other system of logic, a notion referred to as "completeness". Conversely, no false assertion has a proof, a notion referred to as "soundness".

The ability to perform such non-trivial manipulations to formats in which theorems and proofs are presented raises the possibility that we may specify formats that allow for other features (that one does not expect from classical proofs). The notion of PCPs emerges from this study. Here we consider verifiers that vary in two senses: (1) The verifiers are probabilistic — they have access to a sequence of unbiased independent coins (i.e., random variables taking on values from the set $\{0, 1\}$); and (2) The verifiers have "oracle" access to the proof. I.e., to read any specific bit of the proof the verifier is allowed direct access to this bit and charged one "query" for this access. (This is in contrast to the classical notion of the Turing machine where all information is stored on tapes and accessing the ith bit takes i units of time and implies access to all the first i bits of the proof.) However, we will restrict the number of random bits that the verifier has access to. We will also restrict the number of queries the verifier is allowed to make. The latter is definitely a restriction on the power of the verifier (classical verifiers accessed every bit of the proof). The former does not enhance the power of the verifier *unless* the verifier is allowed to err. So we will allow the verifier to err and consider the question: It must be stressed at this point that we require the error probability is bounded away from 1 for *every* false assertion and *every* supporting argument. (It would not make any sense, given the motivation above to assume some random distribution over theorems

and proofs, and this is not being done.) What is the tradeoff between the query complexity and the error incurred by the verifier?

Theoretical computer scientists started to examine this tradeoff starting 1990 and have made some remarkable progress to date. We review this history below. (We remark that this is just a history of results; the notion of a probabilistically checkable proof itself evolved slowly over a long sequence of works [17, 6, 9, 15, 4, 14, 3], but we will not describe the evolution of this notion here.) Results constructing PCP verifiers typically restrict the number of random bits to be logarithmic in the size of the probabilistically checkable proof. Note that this is an absolute minimum limit, or else a verifier making few queries does not have a positive probability of accessing most of the bits of the proof. They then asked the question: How small the can the PCP be (relative to the classical proof) and how many bits needed to be queried? The first sequence of results [5, 4, 14] quickly established that the number of queries could be exponentially smaller than the length of the proof (e.g., in a proof of length n, the number of queries may be as small as say $\log^2 n$), while getting nearly polynomial sized proofs (in fact, [4] obtained nearly linear sized PCPs.) The second short sequence [3, 2] established what is now referred to as "The PCP Theorem" which showed that the number of bits queried could be reduced to an absolute constant(!) independent of the length of the theorem or the proof (given just the length of the proof), with PCPs of length just a polynomial in the classical proof. This immediately raised the question: What is this universal constant — the number of queries that suffices to verify proofs probabilistically. It turns out there is yet another tradeoff hidden here. It is always possible to reduce the number of queries to three bits, if the verifier is allowed to err with probability very close to (but bounded away from) one. So to examine this question, one needs to fix the error probability. So, say we insist that arguments for incorrect assertions are accepted with probability (close to) half, while proofs of valid theorems are accepted with probability one. In such a case, the number of queries made by the verifier of [2] has been estimated at around 10^6 bits - not a dramatically small constant, though a constant all right! The third phase in the construction of PCPs [8, 7] attempted to reduce this constant and culminated in yet another surprise. Hastad [18] shows that the query complexity could be essentially reduced to just *three* bits to get the above error probabilities. Subsequent work in this area has focussed on the question of the size of the PCP relative to the size of the classical proofs and shown that these could be reduced to extremely blow-ups. (Classical proofs of length n are converted to PCPs of length $n \cdot (\log n)^{O(1)}$ in the work of Dinur [12].)

A somewhat orthogonal goal of research in PCPs has been to find simple reasons why proofs ought to be probabilistically checkable. Unfortunately, much of the above results did not help in this regard. The results from the first sequence achieved the effect by a relatively straightforward but striking algebraic transformation (by encoding information into values of algebraic functions over finite fields). Later results built on this style of reasoning but got even more complex (see e.g., [25, Page 12] for a look at the ingredients needed to get the PCP theorem of [18]). Recently, Dinur and Reingold [13] proposed a novel, if somewhat ambitious, iterative approach to constructing PCPs, which was radically different than prior work. While the idea was appealing, the specific implementation was still hard, and did not lead to a satisfactory alternative construction of PCPs. Subsequently, Dinur [12] finally made remarkable progress on this question deriving the right ingredients to give a dramatically simple proof of the PCP theorem.

This work of Dinur is the focus of the rest of this article. Our intent, however, is not to give Dinur's proof of the PCP theorem. This is already done quite satisfactorily in her work [12]. Instead we will try to outline her approach and provide context to the steps taken in Dinur which may provide further insight into her work (and highlight the novelty of the approach as well as the new technical

ingredients developed in her work). The hope is that a reader, after reading this article, would be motivated to read the original work, and upon doing so, appreciate the developments in her paper.

In what follows, we will start by formally describing PCPs and the PCP theorem. Readers uncomfortable with boring formalisms could skip this section. Next we describe a duality between PCP verifiers and "approximations to combinatorial optimization problems". We will use this duality to switch our language from the "logical" theme of theorems and proofs, to a more "combinatorial" theme. (A reader who chooses to skip this section would be lost thereafter.) In Section 4 we then describe the high-level approach in Dinur's paper and contrast it with the earlier approaches. Dinur's approach repeatedly applies two transformations to a "current verifier", starting from a classical (non-probabilistic) verifier of proofs. The end result is a probabilisitic verifier of proofs. In Sections 5 and 6 we describe the two transformations in greater detail providing background on these (in particular, we describe some simpler transformations one may consider, and why they don't work).

2 Definitions and formal statement of results

We start by recalling the notion of a classical verifier and introducing some notation.

First some general notation for the paper. Below $\mathbb{R}$ will denote the reals, $\mathbb{Z}$ the set of all integers, and $\mathbb{Z}^+$ the set of positive integers. For $x \in \mathbb{R}$, we let $\lfloor x \rfloor$ denote the largest integer less than or equal to x. For $x \in \mathbb{R}$, let $\log x$ denote the quantity $\lceil \log_2 x \rceil$ where $\log_2$ denotes the logarithm of x to base 2.

By $\{0,1\}^*$ we denote the set of all finite length binary sequences. (We refer to such sequences as strings.) For a string $x \in \{0,1\}^*$, let $|x|$ denote its length. For random variable X taking on values in domain D and event $E : D \in \{\text{true}, \text{false}\}$, we let $\Pr_X[E(X)]$ denote the probability of the event E over the random choice of X. We often use the shorthand "$f(n)$" to denote the function $n \mapsto f(n)$. (In particular, it will be common to use "n" to denote the argument of the function, without explicitly specifying so.) Examples include the functions n^2, $\log n$ etc.

Later in the writeup we will need to resort to some "graph theory". By a graph we refer to symmetric pairwise relationships on some finite set. Formally a graph G is given by a pair (V, E) with V being a finite set and $E \subset V \times V$ is a symmetric relationship. If $(u, v) \in E$, then we refer to v as being adjacent to u, or being a neighbor of u. The number of vertices adjacent to u is called the degree of u. We say a graph has degree D if every vertex has degree D. A walk in a graph is a finite sequence of vertices $v_0, \ldots, v_\ell$ such that v_{i-1} and v_i are adjacent for every $i \in \{1, \ldots, \ell\}$. The distance between u and v is the length ℓ of the shortest walk $v_0, \ldots, v_\ell$ satisfying $v_0 = u$ and $v_\ell = v$.

We now move to notions related to proof verification. A *verifier* is a polynomial time algorithm computing a function $V : \{0,1\}^* \times \{0,1\}^* \to \{0,1\}$, with the association that $V(T, \Pi) = 1$ implies that the assertion T is a theorem with Π being a proof. (Recall that a function is said to be polynomial time computable if there exists an algorithm running in time bounded by a fixed polynomial in the total length of its inputs to compute the function.) Given a polynomial $p : \mathbb{Z}^+ \to \mathbb{Z}^+$ and verifier V, let $L_{V,p}$ denote the set of theorems with "short" proofs of length at most $p(n)$. I.e., $L_{V,p} = \{T \in \{0,1\}^* | \exists \Pi \in \{0,1\}^{p(|T|)} \text{ s.t. } \mathrm{V}(\mathrm{T}, \Pi) = 1\}$. The class NP is the set of all such sets $\{L_{V,p} | V \text{ is a verifier and } p \text{ is a polynomial }\}$.

As mentioned earlier, we are going to enhance classical algorithms by endowing them with access to random strings and oracles. We will denote random strings just like other strings. A oracle will just be a function $O : \mathcal{Q} \to \mathcal{A}$ where $\mathcal{Q}$ is a countable set and $\mathcal{A}$ is finite. The most common version is with $\mathcal{Q} = \mathbb{Z}^+$ and $\mathcal{A} = \{0,1\}$. Algorithms are allowed to compute various queries $q_1, \ldots, q_t$ and obtain answers $O[q_1], \ldots, O[q_t]$ to the queries. The number of queries made (t) is termed the query complexity of the algorithm. Thus the computation of a probabilistic oracle algorithm A on input x, random string $R \in \{0,1\}^*$ and access to oracle O will be denoted $A^O(x;R)$. Notice that we will always be interested in the distribution of this random variable $A^O(x;R)$ when R is chosen uniformly from set $\{0,1\}^\ell$ (while x and O will be fixed). With this notation in hand we are ready to define PCP verifiers and the complexity class PCP.

Definition 1 *For functions $r, q : \mathbb{Z}^+ \to \mathbb{Z}^+$ an (r, q, a)-restricted PCP verifier is a probabilistic oracle algorithm V that on input $x \in \{0,1\}^n$, expects a random string $R \in \{0,1\}^{r(n)}$ and queries an oracle $\Pi : \mathbb{Z}^+ \to \{0,1\}^{a(n)}$ at most $q(n)$ times and computes a "Boolean verdict" $V^\Pi(x;R) \in \{0,1\}$.*

Definition 2 *For functions $c, s : \mathbb{Z}^+ \to [0,1]$ with $0 \leq s(n) < c(n) \leq 1$ for every $n \in \mathbb{Z}^+$, we say that an (r, q, a)-restricted PCP verifier V accepts a set $L \subseteq \{0,1\}^*$ with completeness c and soundness s if for every $x \in \{0,1\}^n$ the following hold:*

Completeness: *If $x \in L$ then there exists a $\Pi : \mathbb{Z}^+ \to \{0,1\}^{a(n)}$ such that $\Pr_R[V^\Pi(x:R) = 1] \geq c(n)$.*

Soundness: *If $x \notin L$ then for every $\Pi : \mathbb{Z}^+ \to \{0,1\}^{a(n)}$ it is the case that $\Pr_R[V^\Pi(x:R) = 1] \leq s(n)$. By $PCP_{c,s}[r, q, a]$ we denote the class of all sets L such that there exists an (r, q, a) restricted PCP verifier accepting L with completeness c and soundness s.*

Throughout this article we will assume that the queries of the PCP verifiers are made "non-adaptively". I.e., the exact location of questions does not depend on the responses to other questions. The responses only affect the accept/reject predicate of the verifier.

As described above the class PCP is significantly over parametrized. These different parameters are useful when describing various (steps in) constructions of PCPs, but for now they are likely to burden the reader. So lets discard a few to derive a simpler collection: All early PCP results were phrased in terms of verifiers that achieved perfect completeness $c(n) = 1$; and soundness $s(n) \leq \frac{1}{2}$. They also fixed $a(n) = 1$ — i.e., oracles responded with one bit per query. Letting $\mathrm{PCP}[r, q]$ denote the class of languages with such restrictions, the early results [5, 4, 14] could be described as showing that there exist polynomials $p_1, p_2 : \mathbb{Z}^+ \to \mathbb{Z}^+$ such that $\mathrm{NP} \subseteq \mathrm{PCP}[p_1(\log n), p_2(\log n)]$. The PCP Theorem, whose new proof we hope to outline later, may now be stated formally as.

Theorem 3 ([3, 2]) *There exist a constant q such that $NP = \cup_{c \in \mathbb{Z}^+} PCP[c \log n, q]$.*

Finally, the state of the art result along these lines is that of Håstad [18], which shows that for every $\epsilon > 0$, $\mathrm{NP} = \cup_{c \in \mathbb{Z}^+} \mathrm{PCP}_{1-\epsilon, \frac{1}{2}+\epsilon}[c \log n, 3]$.

One aspect we do not dwell on explicitly is the size of the "new proof". It is easy to convert an (r, q)-PCP verifier into one that runs in time $2^{r(n)} \times 2^{q(n)}$, whose queries are always in the range $\{1, \ldots, 2^{r(n)+q(n)}\}$. In other words one can assume "w.l.o.g." that the proof is a string of size at most $2^{r(n)+q(n)}$. So in particular if the randomness and query complexity are bounded by $O(\log n)$, then the PCP proofs are still polynomial sized, and so we won't worry about the size explicitly.

3 Optimization and approximation

As alluded to earlier, one of the principal motivations for studying proofs from a computational perspective is that they shed light on the tractability of many computational tasks. For instance, the theory of NP-completeness says that a vast collection of combinatorial optimization problems, such as the "Travelling Salesman Problem" (TSP), (given an $n \times n$ matrix of distances between n cities, find the smallest tour that visits all n cities), the "Independent Set Problem" (given a set of n elements and a list of incompatible pairs, find the largest subcollection that consists no pair of incompatible elements) or the "Knapsack Problem" (given the weights and values of n elements and a bound C, find a subset of elements whose weight sums to less than C, while maximizing the sum of their values), the theory of NP-completeness shows that finding optimal solutions is as hard as finding "proofs" for generic theorems. Formally, given any assertion and a bound on the length B of its proof, one can construct an instance of the NP complete problem, say TSP, and an integer B', such that any solution to the TSP of length at most B' implies that the given assertion is true and has a proof of length at most B. Thus an algorithm to find optimal tours is a generic theorem prover which needs to only know the length of the proof to generate the proof. Indeed much of the strength for the belief that P $\neq$ NP may be attributed to the belief that we don't expect theorem-proving to be automated.

In this context it may make sense that a "probabilistic" notion of checking proofs may lead to some further insight on the complexity of solving combinatorial optimization problems. This guess turns out be true, and it turns out that the existence of PCP verifiers implies that for many of these optimization problems finding "nearly optimal" solutions is as hard as finding optimal solutions. This connection was first made by Feige et al. [14], who showed that the PCP theorem (then still a conjecture) would imply that the independent set size could not be approximated to within constant factors. Subsequently, many other optimization problems were shown to be hard to approximate using the PCP theorem (cf. [2, 21]).

Remarkably, Irit Dinur's proof uses a "folklore" reverse connection which shows that "reductions" showing hardness of approximating some optimization problems can a folklore one) that shows that "hardness of approximating yield PCP verifiers. We describe the optimization problem, used in her proof next, and then explain why the inapproximability of this problem yields a PCP verifier next.

Definition 4 (Constraint Satisfaction Problem (Max k-CSP-Σ**))** *For a finite set Σ and integer k, an input to the problem* Max k-CSP-Σ *consists of m constraints $C_1, \ldots, C_m$ on n variables $X_1, \ldots, X_n$, where a constraint C_j consists of a function $f_j : \Sigma^k \to \{0,1\}$ and k indices $i_1(j), \ldots, i_k(j) \in \{1, \ldots, n\}$. An assignment $\langle X_1, \ldots, X_n \rangle \leftarrow \langle a_1, \ldots, a_n \rangle \in \Sigma^n$* satisfies *the constraint C_j if $f_j(\alpha_1, \ldots, \alpha_k) = 1$ where $\alpha_\ell = a_{i_\ell(j)}$. The goal is to compute an assignment, given $C_1, \ldots, C_m$, that maximizes the number of constraints that are satisfied.*

Since Max k-CSP-Σ occupies a central role in this article, let us introduce some notation that will be useful later. We often use ϕ to denote instances of Max k-CSP-Σ and $\vec{a} \in \Sigma^n$ to denote an assignment to the n variables. For a pair $\phi, \vec{a}$ as above, we use the notation $\phi(\vec{a})$ to denote the number of constraints of ϕ satisfied by $\vec{a}$. An instance ϕ of Max k-CSP-Σ is said to be *satisfiable* if there exists an assignment satisfying all constraints. The unsatisfiability of the instance ϕ, denoted UNSAT(ϕ), is the quantity $\min_{\vec{a}}\{1 - \phi(\vec{a})/m\}$ i.e., the minimum fraction of constraints left unsatisfied by any assignment.

Max k-CSP-Σ problems arise naturally in the theory of NP completeness. The classical 3SAT problem is easily captured as an instance of Max3-CSP-$\{0,1\}$ where the goal is to distinguish satisfiable instances from instances that are not satisfiable. Similarly, the classical 3-coloring problem (given a graph on n vertices, determine if it is possible to color the vertices with three colors $\{R,G,B\}$ such that no edge of the graph is monochromatic), can also be expressed as an instance of Max2-CSP-$\{R,G,B\}$. Indeed it is a classical result [16] that for every $k \geq 2$ and every Σ with $|\Sigma| \geq 2$, it is NP-hard to find optimal solutions to Max k-CSP-Σ.

However, the classical result does not say anything about solving the problem near-optimally. In particular, the state of knowledge prior to the PCP theorem allowed for the possibility that some polynomial algorithm could, on input ϕ for which there exists an assignment satisfying t out of m constraints, always produce an assignment satisfying $t(1-o(1))$ constraints! Indeed this may be a good point to introduce the notion of an approximation algorithm.

Definition 5 *For $\alpha \geq 1$ An algorithm A that takes as input an instance ϕ and in polynomial time outputs an assignments $\vec{a} \in \Sigma^n$ such that $\phi(\vec{a}) \geq \phi(\vec{a}')/\alpha$ for every other assignment $\vec{a}' \in \Sigma^n$ is called an α-approximation algorithm for* Max k-CSP-Σ.

The PCP theorem rules out the possibility of α-approximation algorithms for Max k-CSP-Σ, unless NP=P. The following proposition gives a weak version of this result.

Proposition 6 *Let q be the constant from Theorem 3. If there is a $(2-\epsilon)$-approximation algorithm for* Max$q$-CSP-$\{0,1\}$ *for any $\epsilon > 0$, then P=NP.*

Proof Sketch: Let A be a $2-\epsilon$ approximation algorithm for Max$q$-CSP-$\{0,1\}$. Let L be a language in NP we wish to decide. Let $V = V_L$ be the $(r(n) = O(\log n), q)$-PCP verifier for this language as guaranteed by Theorem 3. Now consider a string $x \in \{0,1\}^n$ for which we wish to know if $x \in L$ or not. Let $\ell \leq 2^{r(n)+q}$ denote the size of the PCP proof that V queries to check membership of $x \in L$. Denote by $X_1, \ldots, X_\ell$ the Boolean variables representing the oracle responses to the verifiers queries. Now for each random string $R \in \{0,1\}^{r(n)}$ create a q-ary constraint C_R as follows: Let $i_1, \ldots, i_q$ be the q queries made by V on input x and random string R. Furthermore, let $f = f(A_1, \ldots, A_q)$ denote the verifier's acceptance predicate on responses A_t to query i_t. Let $C_R = (f, (i_1, \ldots, i_q))$ be the Rth constraint. Let $\phi = \langle C_R \rangle_{R \in \{0,1\}^{r(n)}}$ be the instance of Max$q$-CSP-$\{0,1\}$ thus obtained.

It is easy to verify that $\mathrm{UNSAT}(\phi) = 1 - \max_\Pi\{\Pr_R[V^\Pi(x;R)]\}$. Thus, if $x \in L$ then ϕ is satisfiable and if $x \notin L$ then $\mathrm{UNSAT}(\phi) \geq \frac{1}{2}$. Now consider running A on ϕ. If $x \in L$, then $A(\phi)$ produces an assignment satisfying $m/(2-\epsilon)$ constraints, where $m = 2^{r(n)}$. On the other hand, if $x \notin L$, no assignment satisfies more than $m/2$ constraints. Thus to decide if $x \in L$, all we need to do is to count the number of assignments satisfied by $A(\phi)$ and accept iff this number is more than $m/2$. Since the transformation of x to ϕ takes only polynomial time, and A runs in polynomial time, this gives a polynomial time algorithm to solve a generic NP language L, thus yielding NP=P. ∎

We use the phrase "inapproximable to within a factor of α" to denote that existence of an α-approximation algorithm would imply $P = NP$.

The above proposition and proof only cover the case of Max k-CSP-Σ for some choice of k and Σ. However standard reductions can then be used to show that Max k-CSP-Σ is inapproximable to

within some constant $\alpha > 1$ for every $k \geq 2$ and every Σ with $|\Sigma| \geq 2$. We will elaborate on this later.

The proof above shows that to show that Max k-CSP-Σ is α-inapproximable, it suffices to produce a reduction of the following form for some NP complete language L: The reduction should map, in polynomial time, an instance $x \in \{0,1\}^n$ to an instance ϕ of Max k-CSP-Σ such that ϕ is satisfiable if $x \in L$ and $\text{UNSAT}(\phi) \geq 1 - \frac{1}{\alpha}$. The following proposition shows that any such reduction implies the PCP theorem.

Proposition 7 *Suppose there is a polynomial time reduction from an NP complete language L to* Max k-CSP-Σ *mapping an instance x to ϕ such that ϕ is satisfiable if $x \in L$, and* $\text{UNSAT}(\phi) \geq \epsilon$ *if $x \notin L$. Then $L \in PCP_{1,1-\epsilon}[O(\log n), k, \log |\Sigma|]$.*

Proof Sketch: The verifier for the assertion "$x \in L$" uses the reduction to produce an instance ϕ of Max k-CSP-Σ. It then expects as proof an oracle Π giving the assignment satisfying ϕ. (So $\Pi[i] = a_i$ where $\vec{a} = \langle a_1, \ldots, a_n \rangle$ is the assignment satisfying ϕ.) To verify the proof, the verifier picks a random constraint C_j of ϕ and verifies it is satisfied by Π. Notice thus that the verifier makes k queries to the proof oracle, getting an element of Σ (which can be encoded by $\log |\Sigma|$ bits) as response. It can also be verified that the verifier accepts with probability one if $x \in L$ and with probability at most $1 - \epsilon$ if $x \notin L$. ∎

Dinur's proof directly produces a reduction showing such a hardness. We state her main theorem below.

Theorem 8 ([12]) *There exists an NP complete language L, $\epsilon > 0$, finite set Σ, and a polynomial time reduction R, mapping instances x to ϕ of* Max2-CSP-Σ *such that ϕ is satisfiable if $x \in L$ and* $\text{UNSAT}(\phi) \geq \epsilon$ *if $x \notin L$.*

Combined with Proposition 7 above, this yields the PCP theorem.

4 Overview of Dinur's approach

Before moving on to describing Dinur's approach to proving the PCP theorem, let us briefly describe the prior approaches. The prior approaches to proving the PCP theorem were typically stated in the "$\text{PCP}_{c,s}[r,q,a]$" notation, but the effective equivalence with Max k-CSP-Σ allows us to interpret them in the CSP notation, and we do so below.

One of the principal issues to focus on is the "Gap" in the unsatisfiability achieved by the reduction. Notice that the reductions we seek achieve a significant gap in the unsatisfiability of the instances achieved when $x \in L$ (which should be 0) and the unsatisfiability when $x \notin L$ (which should be lower bounded by some absolute constant $\epsilon > 0$). We refer to this quantity as the "Gap" of the reduction.

Previous approaches were very careful to maintain large gaps in reductions. Since it was unclear how to create a direct reduction from some NP complete language L to Max k-CSP-Σ for finite k and Σ with a positive gap, the prior approaches considered allowing k and Σ to grow with $n = |x|$. The results of Babai et al. [5, 4] and Feige et al. [14] used algebraic techniques (representing

information as coefficients of multivariate polynomials and encoding them by their evaluations) to get reductions from any NP-complete language L to Max$k(n)$-CSP-$\Sigma(n)$ where $k(n), \log|\Sigma(n)| \approx (\log n)^{O(1)}$. Arora and Safra [3], observed an asymmetry in the behavior of the two parameters $k(n)$ and $\log|\Sigma(n)|$ and in particular observed that one could interpret existing PCP constructions as techniques that reduce Max k-CSP-$\Gamma(n)$ to Max$O(k)csp\Sigma(n)$ where $|\Sigma(n)| \ll |\Gamma(n)|$.

This motivated the search for new PCPs which maintained k to be some absolute constant, while allowing $\Sigma(n)$ to grow. Arora et al. [2] produced two such reductions, one of which reduced Max k-CSP-$\Gamma(n)$ to Max$O(k)$-CSP-$\Sigma(n)$ with $\log|\Sigma(n)| \approx (\log\log|\Gamma(n)|)^3$, and another reduction reducing Max k-CSP-$\Sigma(n)$ to Max$O(k)$-CSP-$\{0,1\}$ (but with the catch that the reduction took time that was at least $\Sigma(n)$, so one couldn't afford to use it on large $\Sigma(n)$). Each one of these reductions reduced the gap by a constant factor, but this was ok since one only needed to apply these reductions a constant number (thrice in [2]) to reduce the alphabet size $\Sigma(n)$ to an absolute constant.

Thus the previous approach could be described as constructing PCPs by "alphabet reduction", subject to "gap preservation". In contrast, Dinur's approach seems to be quite the opposite. In her approach, she starts with a reduction from the NP complete language L to Max k-CSP-Σ which has minimal gap (producing only UNSAT$(\phi) \geq 1/m$ when $x \notin L$), but where k and Σ are finite. She then applies a sequence of iterations that ensure "gap amplification" while "preserving alphabet size". The following lemma, from which the main theorem follows easily describes the properties of these iterations.

Lemma 9 (Main Lemma) *There exists a finite set Σ, a positive constant $\epsilon > 0$ and a linear time reduction*[1] *T transforming instances of* Max 2-CSP-Σ *to instances of the same problem such that*

Completeness *ϕ is satisfiable $\Rightarrow$ $T(\phi)$ is satisfiable.*

Soundness UNSAT$(\phi) \geq \min\{2\text{UNSAT}(\phi), \epsilon\}$.

The reduction above is totally novel in the PCP literature and already finds other applications (other than providing alternate proofs of the PCP theorem) in Dinur's paper (see [12, Section 7]). Indeed a few iterations (logarithmically many) of the transformation above amplifies the gap of any reduction from tiny amounts to an absolute constant, and thus yields the PCP theorem. The following proof argues this formally.

Proof of Theorem 8: Given an NP complete language L and a string $x \in \{0,1\}^n$ for which we wish to decide membership, we first transform it to an instance ϕ_0 of Max 2-CSP-Σ such ϕ_0 is satisfiable if and only if $x \in L$. (Notice such reductions, with effectively trivial gap, are classical.) Let m denote the number of constraints of ϕ_0. Now iterate the transformation T from Lemma 9 $\ell = \log m$ times, and let $\phi_i = T(\phi_{i-1})$. We claim that the reduction that maps x to ϕ_ℓ has the properties claimed in the theorem.

First note that if $x \in L$, then ϕ_i is satisfiable for every $i \in \{0, \ldots, \ell\}$ satisfying the "completeness" condition.

[1] I.e., there exist absolute constants c, d such that $T(\phi)$ takes time at most $c|\phi| + d$ to compute. In particular, this implies that $|T(\phi)| \leq c|\phi| + d$.

Next note that if $x \notin L$, then $\text{UNSAT}(\phi_0) \geq 1/m$ (since ϕ_0 is not satisfiable). By induction (using the Soundness condition in Lemma 9) we can now see that $\text{UNSAT}(\phi_i) \geq \min\{\frac{2^i}{m}, \epsilon\}$. Thus, since $2^\ell \geq m$, we have $\text{UNSAT}(\phi_\ell) \geq \epsilon$.

Finally, we need to argue that the entire reduction takes polynomial time. To do this it suffices to argue that the size of the instance ϕ_ℓ is only polynomially larger than x. (The total running time is then bounded by the time take to produce ϕ_0 plus at most $\log m$ times some linear function in $|\phi_\ell|$.) To argue this we use (in fact, need!) the fact that T is a linear time reduction and so $|T(\phi)| \leq c|\phi| + d$. For simplicitly, assume $d = 0$. Then by induction, we see that $|\phi_\ell| \leq c^\ell \cdot |\phi_0| \leq O(m^{\log_2 c}) \cdot |\phi_0| \leq (|\phi_0|)^{O(1)} \leq (|x|)^{O(1)}$ as required. ∎

Thus our focus now shifts to Lemma 9 and we start to peek into its proof. Dinur's proves this Lemma by combining two counteracting reductions. The first reduction amplifies the gap by increasing the alphabet size. Since this is the main novelty in Dinur's reduction, we will defer its proof to the end. The second reduction is now in the classical style, which reduces the gap (somewhat), while reducing the alphabet size. While it is clear that both reductions are opposing in direction, the level of detail used above leaves it unclear as to what would happen if the two reductions were applied in sequence. Would this increase the gap or reduce it? Would it increase the alphabet size or reduce it (or preserve it)?

Part of the insight behind Dinur's approach is the observation that both these reductions are especially strong. The first allows gap amplification by any amount, subject to a sufficiently large explosion in the alphabet size. The second reduction can reduce any alphabet to a fixed small alphabet, while paying a fixed price in terms of the gap. These terms are articulated in the assertions below.

Lemma 10 *For every constant $c < \infty$ and finite set Σ there exist constant $\epsilon_1 > 0$, finite set Γ and a linear time reduction T_1 from* Max 2-CSP-Σ *to* Max 2-CSP-Γ *such that:*

Completeness *ϕ is satisfiable $\Rightarrow$ $T_1(\phi)$ is satisfiable.*

Soundness $\text{UNSAT}(T_1(\phi)) \geq \min\{c \cdot \text{UNSAT}(\phi), \epsilon_1\}$.

(In other words, one can pick any amount to amplify by, and the reduction finds an appropriate alphabet Γ to reduce to.)

Lemma 11 *There exists a constant $\epsilon_2 > 0$, a finite set Σ such that for every finite Γ, there exists a linear time reduction T_2 mapping* max 2*cspg to* max 2*csps such that:*

Completeness *ϕ is satisfiable $\Rightarrow$ $T_2(\phi)$ is satisfiable.*

Soundness $\text{UNSAT}(T_2(\phi)) \geq \epsilon_2 \cdot \text{UNSAT}(\phi)$.

Notice that ϵ_2 above — the loss in the gap — is independent of alphabet size. We will elaborate more on this in the next section.

We defer the proofs of the two lemmas to the ensuing sections, but now show how the main lemma follows from the above two.

Proof of Lemma 9: Let Σ, ϵ_2 be as in Lemma 11. Let $c = 2 \cdot \epsilon_2$. Invoking Lemma 10 for this choice of c and Σ, let ϵ_1, Γ and T_2 be as given by Lemma 11. Now invoke Lemma 11 for this choice of Γ and let T_2 be the reduction so obtained.

We claim Lemma 9 holds for Σ, $\epsilon = \epsilon_1 \cdot \epsilon_2$ and T being the composition of T_2 with T_1.

It is clear that $T_2(T_1(\cdot))$ maps instances of Max 2-CSP-Σ to instances of the same problem. Since both T_1 and T_2 are linear time reductions, it follows that so is T. Also since both preserve satisfiability, so does their composition. Finally the unsatisfiability of $T(\phi)$ may be lower bounded as follows.

$$\mathrm{UNSAT}(T_2(T_1(\phi))) \geq \epsilon_2 \cdot \mathrm{UNSAT}(T_1(\phi)) \geq \min\{\epsilon_2 \cdot c \cdot \mathrm{UNSAT}(\phi), \epsilon_2 \cdot \epsilon_1\} = \min\{2 \cdot \mathrm{UNSAT}(\phi), \epsilon\}.$$

This concludes the proof of Lemma 9. ∎

In the following sections we comment on the proofs of Lemmas 10 and 11. Since the former is the more novel element, we defer discussion about it to the end. We start with Lemma 11.

5 Alphabet Reduction and Error Correcting Codes

In order to motivate the strength of Lemmas 11 and 10 we first describe some of the more elementary operations one can use to manipulate the parameters k and Σ. The results in the following proposition are by now either considered "basic" or at least "standard" in the context of approximation preserving reductions. The reader is strongly encouraged to think about each individually before reading the proof sketch, so as to gain some intuition into the assertions and their proofs.

Proposition 12 *Fix integers ℓ, k. There exist linear-time, satisfiability preserving reductions A_1, A_2, and A_3 such that*

1. *A_1 reduces* Max$k$-CSP-$\{0,1\}^\ell$ *to* Max$(k \cdot \ell)$-CSP-$\{0,1\}$ *with* $\mathrm{UNSAT}(A_1(\phi)) = \mathrm{UNSAT}(\phi)$.
2. *A_2 reduces* Max$k$-CSP-$\{0,1\}$ *to* Max3-CSP-$\{0,1\}$ *with* $\mathrm{UNSAT}(A_2(\phi)) \geq \frac{1}{2^{k+2}}\mathrm{UNSAT}(\phi)$.
3. *A_3 reduces* Max$k$-CSP-$\{0,1\}$ *to* Max2-CSP-$\{0,1\}^k$ *with* $\mathrm{UNSAT}(A_2(\phi)) \geq \frac{1}{k}\mathrm{UNSAT}(\phi)$.

Proof Sketch: We consider the items in sequence.

1. For the first part, given a Max$k$-CSP-$\{0,1\}^\ell$ instance ϕ with constraints $C_1, \ldots, C_m$ on variables $X_1, \ldots, X_n$ taking values in $\{0,1\}^\ell$, we "encode" each variable X_i by a collection of ℓ Boolean variables $Y_{i,j}$, $j \in \{1, \ldots, \ell\}$, with the association that an assignment $\vec{a} = \langle a_1, \ldots, a_\ell \rangle \in \{0,1\}^\ell$ to X_i corresponds to the assignments $Y_{i,j} \leftarrow a_j$. A constraint $C_j = f(X_{i_1}, \ldots, X_{i_k})$ can now be naturally represented as a constraint $C'_j = f'(Y_{i_1,1}, \ldots, Y_{i_1,\ell}, \ldots, Y_{i_k,1}, .$ where f' is satisfied by an assignments to the $Y_{i,j}$'s if and only if the corresponding assignment to the X_i's satisfies f. It is easy to verify that the assignments to X_i's are in 1-to-1 correspondence with the assignments to $Y_{i,j}$'s with corresponding assignments satisfying exactly the same number of constraints. This yield the reduction A_1.

   (The important aspect to note in this reduction is that its performance degrades with ℓ. Indeed this is one of the principal effects we will aim to remedy later.)

2. For the second part, we hint that this is essentially similar to the classical reduction from "SAT" to "3SAT", whose approximability properties were clarified in [23]. In this reduction, when transforming an instance ϕ with constraints $C_1, \ldots, C_m$ on variables $X_1, \ldots, X_n$, one retains all the original variables, and adds for each constraint $C_j = f(X_{i_1}, \ldots, X_{i_k})$ a collection of $K \approx 2^k$ "auxiliary" variables $Y_{j,1}, \ldots, Y_{j,K}$ and introduce K ternary constraints $C'_{j,1}, \ldots, C'_{j,K}$ on variables $(X_{i_1}, \ldots, X_{i_k}, Y_{j,1}, \ldots, Y_{j,K})$ such that an assignment to $(X_{i_1}, \ldots, X_{i_k}) \leftarrow \vec{a}$ satisfies C_j if and only if there exists an assignment $(Y_{j,1}, \ldots, Y_{j,K}) \leftarrow \vec{b}$ such that all the constraints $C'_{j,1}, \ldots, C'_{j,K}$ are satisfied by the assignment $(\vec{a}, \vec{b})$. It can be seen that this reduction has the right properties.

3. This reduction, though also simple, is more "recent" than others, having been first brought out by the work of Fortnow et al. [15]. Here, the idea is to lump together k bit strings queried in various constraints as new single variables, but then to check their consistency against the older single bit assignments. Formally, given k-ary constraints $C_1, \ldots, C_m$ on Boolean variables $X_1, \ldots, X_n$, we create an instance with km constraints $\{C_{j,\ell}\}$ on $n + m$ variables $X'_1, \ldots, X'_n, Y_1, \ldots, Y_m$. If the constraint $C_j = f(X_{i_1}, \ldots, X_{i_k})$, then the constraint $C_{j,\ell}$ applies to variables Y_j and X'_{i_ℓ} and verifies that the k-bit assignment to $f(Y_j) = 1$, that $X'_{i_\ell} \in \{0^k, 0^{k-1}1\}$, and that the last bit of X'_{i_ℓ} equals the ℓth bit of Y_j. It is easy to see that assignments to the X' variables can be interpreted as assignments to the X variables and that the constraints $C'_{j,1}, \ldots, C'_{j,k}$ are all satisfied only if the corresponding assignment to X_i's satisfy C_j, which suffices to conclude that this reduction has the desired property.

∎

To summarize, Proposition 12 suggests a number of obvious reductions between constraint satisfaction problems. The upshot is that large gaps are hard to achieve when k and $|\Sigma|$ are small. But as it turns out the two parameters, k and $\log|\Sigma|$ are not totally similar in behavior. On the one hand, one can tradeoff Σ for a smaller alphabet, by increasing the number of queries. But reversing this tradeoff does not seem to be as obvious (and more involved results show that we do have to lose something in the unsatisfiability).

Returning to our goal of Lemma 11, of reducing a large alphabet Γ to some small fixed alphabet Σ, we see we could do this, if we were allowed to increase the number of queries (but we have to keep this fixed to 2), or allow the unsatisfiability of the reduced instance to be much smaller than (such as say $1/|\Gamma|$ times) the unsatisfiability of the source instance. But we wish to do better and lose only a fixed constant.

Turns out the prior work on PCPs, in particular [2, Section 6], addresses precisely this issue (though it was not conceived to be utilized as many times as in the current proof). The steps in the reduction resemble the classical one (Proposition 12, Part 1) however each step is significantly different. Given an instance ϕ of max 2-CSP-Γ with constraints $C_1, \ldots, C_m$ on variables $X_1, \ldots, X_n$, we first produce an instance of max 3-CSP-$\{0,1\}$ with the following steps:

1. First we "encode" each variable X_i taking values in Γ with a collection of Boolean variables $X'_{i,1}, \ldots, X'_{i,K}$ (for some large constant K depending only on $|\Gamma|$). The classical reduction did so by representing elements of Γ as binary strings and then using the new variables to represent these binary strings (see the proof of Proposition 12, Part 1). Unfortunately this representation is not robust, and loses $1/\log|\Gamma|$ factor in the gap simply due to the fact that

two different elements of Γ may differ in only one bit in their respective binary encodings. The reduction used to prove Lemma 11 in [12] gets around this loss by representing elements of Γ in an "error-correcting code": Specifically, we find a collection S of Γ strings in $\{0,1\}^\ell$ for an appropriate integer ℓ so that every pair of strings differ in, say, at least $\ell/10$ coordinates. Such codes are well known to exist, though for our purposes it is more convenient to work with special codes.

2. Next, for a constraint, $C_j = f(X_{i_1}, X_{i_2})$, we introduce a new collection of variables $\{Y_{j,t}\}_t$ and a collection of constraints $\{C'_{j,t}\}$ on the variables $\{X'_{i_1,1}, \ldots, X'_{i_1,K}\}, \{X'_{i_2,1}, \ldots, X'_{i_2,K}\}, \{Y_{j,t}\}_t$. We won't be able to describe these constraints here, but they "enforce" two conditions: (1) They enforce that the variables $X'_{i_1,*}$, $X'_{i_2,*}$ are close encoding of some strings in the code S (e.g., changing fewer than $\ell/5$ variables $X'_{i_1,*}$ yields a string in S). (2) They enforce that the closest members of S correspond to assignments to X_{i_1} and X_{i_2} that satisfy C_j. The special aspect of the new constraints is that even though we have an enormous number of these constraints (growing with $|\Gamma|$), violating either of the conditions (1) or (2) would lead to a constant (say $3\epsilon_2$) fraction of the constraints $\{C'_{j,t}\}_t$ being violated (whereas classical reductions only violated a single constraint, when an original constraint was unsatisfied).

   Finding the right code S, the number of auxiliary variables $Y_{j,*}$ and the right collection of constraints $C'_{j,*}$ may be dismissed as a mere a "finite" search problem, if only we could prove that they exist. Unfortunately, the only proofs that we know that such structures exist, is the constructive one. And the constructive proof essentially amounts to building a "finite" PCP-like object (where failure to satisfy some conditions are "visible" to many local checks). Fortunately, these PCPs can afford to be much larger than the "polynomial sized" PCPs we seek, and their constructions are significantly simpler. Dinur presents a very compact such construction (see [12, Section 6]) while earlier constructions (e.g., [2, Section 6]) while being longer are still quite simple and natural. We remark that while the problem is a very combinatorial one, the construction of these gadgets and their analysis does rely on "algebraic" results over finite fields ([2]) or Harmonic analysis over the Boolean cube ([12]).

Once one has such a reduction from Max 2-CSP-Γ to Max3-CSP-$\{0,1\}$ one can apply a standard reduction (Proposition 12, Part 3) to now reduce the problem further to max 2-CSP-$\{0,1\}^3$ yielding Lemma 11 for $\Sigma = \{0,1\}^3$. The reader may look at [12, Section 6] for details.

6 Gap amplification

We now move to the technical centerpiece of Dinur's proof of the PCP theorem. Before getting into the specifics of this problem, we first describe the context of the result and its proof.

6.1 Background: Recycling Randomness

The underlying problem here, of amplifying gaps, plays a major role in the developing theory of "randomized computation". Since every essentially randomized algorithm errs with some positive probability, a natural question is to investigate whether this error could be reduced.

For instance, consider one of the classical (randomized) algorithms to determine if an n-bit integer is a prime. The early algorithms (cf. [22]) had the property that they would always declare prime

inputs to be "prime", but for any composite input they may declare it also to be "prime" with probability half. The classical algorithm would need an n-bit long random string to perform this test. Now, suppose we wish to reduce this error probability (of concluding that composites may be "prime") to say 1/128, one only needs to run the basic algorithm 7 times and declare a number to be prime only if every one of the seven iterations declared it to be prime. One of the drawbacks of this approach is that this process costs seven times the original cost in terms of randomness, as well as running time. While the latter may be an affordable cost (esp. for settings other than primality testing where no polynomial time deterministic algorithm is known), however, the increasing cost of randomness may prove less affordable. (Unlike the case of processor speeds in computers which under the empirically observed "Moore's Law" keep doubling every three years, physical generation of pure randomness does not seem to be getting easier over the years.) In view of this, one may ask if there is a more "randomness-efficient" way to get the error probability down to 1/128 without expending $7n$ random bits?

This task has been studied extensively under the label of "recycling randomness" [1, 11, 19] in the CS literature, which shows that it suffices to use something like $n + ck$ bits, for some absolute constant c, to reduce the error to 2^{-k} (though the cost in terms of running time remains a multiplicative factor of k). The most common technique for such "random-efficient" amplification, is to repeat the randomized algorithm with related randomness. More formally, suppose $A(x; R)$ denotes the computation of a randomized algorithm to determine some property of x (e.g., $A(x) = 1$ if and only if x is a prime integer). The standard amplification constructs a new algorithm $A'(x; R')$ where $R' = (R_1, \ldots, R_k)$ is a collection of k independent random strings from $\{0,1\}^n$ and $A'(x; R') = 1$ if and only $A(x; R_1) = \cdots = A(x; R_k) = 0$. Now, given that each invocation $A(x; R_i)$ only "leaks" one bit of information about R_i, using independent random coins is completely inessential for this process. Indeed it is easy to subsets $S \subseteq \{\{0,1\}^n\}^k$ of cardinality only $2^{(O(n+k)}$ such the performance of A' where R' is chosen uniformly from S is almost as good as when drawn from the entire universe of cardinality 2^{nk}. The computational bottleneck here is to produce such a distribution/set S efficiently.

One popular approach to producing such a set efficiently uses the technique of "random walks" on "expander graphs". Here we create a graph G whose vertices are the space of random strings of A (i.e., $\{0,1\}^n$) with the property that each vertex of G is adjacent to a fixed number, D, of other vertices in G. For the application of recycling randomness it will be important that one can enumerate in time polynomial in n all the neighbors of any given vertex $R \in \{0,1\}^n$, though for the purpose of the PCP gap amplification it will suffice to be able to compute this in time $2^{O(n)}$. The "random walk" technique to recycling randomness produces $R' = (R_1, \ldots, R_k)$ by first picking $R_1 \in \{0,1\}^n$ uniformly at random, and then picking R_2 to be a random neighbor of R_1, and R_3 to be a random neighbor of R_2 and so on. In other words R' is generated by taking a "random walk" on G.

To understand the randomness implications of this process, we first note that this process takes $n + k \log D$ bits of randomness. So it is efficient if D is small. On the other hand the amplification property relates to structural properties of the graph. For instance, the reader can see that it wouldn't help if the graph had no edges, or were just a collection of $2^n/(D+1)$ disconnected complete graphs of size $D+1$ each! Indeed for the amplification to work well, the graph needs to be an extremely well connected graph, or an "expander" as defined next.

Definition 13 *For a graph $G = (V, E)$ and subset $S \subseteq V$, let $E_S = \{(u, v) \in E$ s.t. $|\{\mathrm{u}, \mathrm{v}\} \cap \mathrm{S}| = 1\}$ denote the set of edges crossing from S to its complement. The expansion of the set S, denoted*

$e(S)$, is the quantity $\left(\frac{|E_S|}{|E|}\right) / \left(\frac{|S|}{|V|}\right)$. G is said to be a (γ, D)-expander if every vertex is adjacent to exactly D other vertices, and every set S with $|S| \leq |V|/2$ has expansion $e(S) \geq \gamma$.

It is by now well-known in the CS literature that if R' if generated by a k-step random walk on a (γ, D)-expander, that the error probability reduces to $2^{-\delta k}$ where δ is a universal constant depending only on γ and D. (This result was first shown in a specific context by Ajtai et al. [1], and then noted for its general applicability in [11, 19].) Furthermore, a rich collection of "explicit" (γ, D)-expanders have been constructed, allowing for widespread application of this result. See [20] for a survey.

6.2 Amplification of PCPs: Naive approaches

We now return to the issue of amplifying the gap in Max 2-CSP-Σ. The naive approach to this problem would be to "iterate" the associated PCP verifier twice. The following proposition describes this operation in the CSP language.

Proposition 14 *There exists a quadratic time satisfiability preserving reduction A_4 reducing* Max 2-CSP *to* Max4-CSP-Σ *such that if* $\text{UNSAT}(\phi) = \epsilon$ *then* $\text{UNSAT}(A_4(\phi)) = 1 - (1-\epsilon)^2$.

We leave it to the reader to verify the above proposition. The main aspect to notice is that the variables of $A_4(\phi)$ are the same as the variables of ϕ, while $A_4(\phi)$ has a constraint C'_{ij} for every pair of constraints C_i, C_j of ϕ where C'_{ij} represents the conjunction of the constraints C_i and C_j.

We move on to the problems with this reduction. First, this reduction takes quadratic time. More significantly the size of the instance $|A_4(\phi)|$ is really quadratic in $|\phi|$ and this is a price we can not afford. (Logarithmically many iterations of this process would blow the instance size up from n to n^n, which completely destroys any hope of using this to construct PCPs.)

Fortunately, this is an aspect that is readily amenable to the "random walk on expanders" technique. Specifically we can consider a better k-fold amplification reduction A_5 reducing Max 2-CSP-Σ to Max($2k$)-CSP-Σ as follows: The variables of $A_5(\phi)$ are the same as the variables of ϕ. Constraints of $A_5(\phi)$ are generated by first picking k constraints of ϕ by performing a k-step random walk on a (γ, D)-expander G with m vertices (so the vertices of G correspond to constraints of ϕ) and then taking the conjunction of all such constraints. The number of constraints now is only $n \cdot D^k$ which is linear in n if k, D are constant. The analysis used in the general setting of recycling randomness can now be used to prove the following proposition.

Proposition 15 *There exists a constant $\delta > 0$ such that for every k, there exists a linear time satisfiability preserving reduction A_5 reducing* Max 2-CSP-Σ *to* Max($2k$) − -CSP-Σ *such that if* $\text{UNSAT}(\phi) = \epsilon$ *then* $\text{UNSAT}(A_4(\phi)) = 1 - (1-\epsilon)^{\delta k}$.

The amplification effects of the above proposition, as well as the time complexity are now as we would like. However there is still one, fatal, flaw with both reductions above. They do not reduce "binary" constraint satisfaction problems to "binary" constraint satisfaction problems. Instead they reduce them to ($2k$)-ary constraint satisfaction problems, which is also of no use in the iterative approach. So we turn to the problem of preserving the "binary" nature of constraints.

6.3 Background: Parallel Repetition

For this section it is convenient to switch to the PCP language. Consider a PCP verifier V that on input x and random string R two queries $q_1(R)$ and $q_2(R)$ to an oracle $\Pi : \mathbb{Z}^+ \to \Sigma$ and accepts if the responses $a = \Pi(q_1(R))$ and $b = \Pi(q_2(R))$ satisfy $f(R, a, b) = 1$ for some fixed predicate f depending on x.

The naive amplification (corresponding to reduction A_4 described earlier) corresponds to the following verifier V': V' picks two random strings R_1, R_2 from the space of the randomness of V and issues queries $q_1(R_1), q_2(R_1), q_1(R_2), q_2(R_2)$ to Π. If the responses are a_1, b_2, a_2, b_2 then V' accepts if $f(R_1, a_1, b_1) = 1$ and $f(R_2, a_2, b_2) = 1$. The acceptance probability of the modified verifier V' (maximized over Π) is the square of the acceptance probability of V (maximized over Π), which is good enough for us. However it makes 4 queries and this is the issue we wish to address in this section.

One natural attempt at reducing the number of queries may be to "combine" queries in some natural way. This is referred to as parallel repetition of PCPs. In the k-fold parallel repetition we consider an new verifier $V^{\|\otimes k}$ that accesses an oracle $\Pi^{\|\otimes k} : (\mathbb{Z}^+)^k \to \Sigma^k$ (with the association that the k coordinates in the domain correspond to k queries to Π, and the k coordinates in the range to the k responses of Π) and functions as follows: $V^{\|\otimes k}$ picks k independent random strings $R_1, \ldots, R_k$ and queries $\Pi^{\|\otimes k}$ with $(q_1(R_1), \ldots, q_1(R_k))$ and $(q_2(R_1), \ldots, q_2(R_2))$. If the responses of $\Pi^{\|\otimes k}$ are $(a_1, \ldots, a_k)$ and $(b_1, \ldots, b_k)$ then $V^{\|\otimes k}$ accepts if $f(R_i, a_i, b_i) = 1$ for every $i \in \{1, \ldots, k\}$.

One may hope that the error in the k-fold parallel repetition goes down exponentially with k. However, any such hopes are dashed by the following example, which gives a choice of (Σ, f, q_1, q_2) such that the error of the k-fold parallel repetition *increases* exponentially with k.

Example: Let V work with $\Sigma = \{0, 1\}$ and the space of random strings R be $\{0, 1\}$. Let $q_i(R) = i + R$ and let $f(0, a, b) = b$, and $f(1, a, b) = 1 - a$. The reader may verify that for every oracle $\Pi : \{1, 2, 3\} \to \{0, 1\}$ the acceptance probability of V is $\frac{1}{2}$. Furthermore there exist $\Pi^{\|\otimes k}$ for which the acceptance probability of $V^{\|\otimes k}$ is $1 - 2^{-k}$.

The example illustrates some of the many problems with naive hopes one may have from parallel repetition. In the face of the above example one may wonder if any amplification is possible at all in this setting. After many works exploring many aspects of this problem, Raz [24] gave a dramatic *positive*. He considers restricted verifiers whose "question" spaces (the image of $q_1(\cdot)$ and $q_2(\cdot)$) are disjoint, and shows that for such verifiers, error does reduce exponentially with the number of iterations, with the base of the exponent depending only on the acceptance probability of the original verifier, and the answer size $|\Sigma|$. Furthermore, there exist reductions reducing any verifier to a restricted verifier only a constant factor in the gap. (The reader may try to see how one such reduction is implied by Proposition 12, Part 3.) Combined these two steps allow us to amplify the gap in PCPs — but now we have lost the "linear time property".

Is it possible to try parallel repetition while recycling randomness? Given the difficulty in analyzing parallel repetition (Raz's proof, while essentially elementary, is already one of the most intricate proofs seen in the PCP setting) the task of combining it with recycling randomness appears forbidding. Remarkably enough Dinur [12] manages to combine the two techniques and achieve the desired gap amplification, and does so with relatively simple proofs. Among other things, Dinur's realization is that even an example such as the above may not defeat the purpose. For the purposes of Lemma 10 it suffices to show that the acceptance probability goes down, provided it was very high to start with; and that in the remaining cases it remains bounded away from 1 (by say, 2^k).

Dealing with cases where the acceptance probability is very high (e.g., greater than $1-\Sigma^{-k}$) turns out be easier than dealing with the other cases. We now describe Dinur's gap amplification.

6.4 Gap amplification

To describe Dinur's gap amplification lemma we switch back to the terminology of CSPs. Her main idea is to consider the graph underlying a Max 2-CSP-Σ instance, and to impose some structure on the graph on it, and then to generate instances of Max2-CSP-Σ^K based on walks of length k on this graph.

We start by describing the graph G_ϕ underlying a $\max 2csps$ instance ϕ. G_ϕ has n vertices correponding to the n variables of ϕ and (i,j) is an edge if there is some constraint (among the m constraints of ϕ) on the pair of variables X_i, X_j. (As an aside, Dinur's analysis relies essentially on the feature that this graph is "undirected" i.e., $(i,j) \in E \Leftrightarrow (j,i) \in E$, which is in sharp contrast to Raz's setting which requires that $(i,j) \in E$ implies that $(i,i') \notin E$ and $(j,j') \notin E$ for any i', j'.)

As a first step, Dinur performs some preprocessing to ensure that G_ϕ is a (γ, D)-expander. If it is not, she reduces, in linear time, the instance ϕ to a different instance $\tilde{\phi}$ of Max 2-CSP-Σ so that $\mathrm{UNSAT}(\tilde{\phi}) \geq \epsilon_3 \cdot \mathrm{UNSAT}(\phi)$. This preprocessing reduction (from ϕ to $\tilde{\phi}$) is achieved by first transforming ϕ to ϕ_1 so that G_{ϕ_1} has bounded degree, which also uses expanders in a technique going back to the work of [23]. Next it transforms ϕ_1 to $\tilde{\phi}$ by imposing a collection of vacuous constraints ϕ_2 (which are always satisfied) such that G_{ϕ_2} is an expander. It may be verified that if G_{ϕ_2} is a $(2\gamma, D/2)$-expander and G_{ϕ_1} has degree $D/2$, then the union of the two graphs yields a (γ, D) expander. If one can amplify the gap of the instance $\tilde{\phi}$ by c/ϵ_3 factor, then the composition of the two steps amplifies the gap of ϕ by a factor c. This (to simplify our notation) below we assume that G_ϕ is an expander.

We now move to the crux of Dinur's amplification. Given ϕ as above, let k correspond to the number of repetitions we intend to attempt. For $u \in V(G_\phi)$, let $B(u,k)$ denote the set of vertices within a distance of at most k from u. Let $K = \max_u\{|B(u,k)|\} \leq \sum_{i=0}^{k} D^i$. Then the new alphabet $\Gamma = \Sigma^K$. The new instance ϕ' will continue to have n variables (same as ϕ), where the new variable X'_u will be viewed as assigning an opinion on its value of the assignment to X_v for every v that is within a distance of k from u in G_ϕ. (Notice that the number of such v's for any fixed u is at most K and so indeed an alphabet of size $|\Sigma|^K$ suffices to represent all these opinions.) We use $X'_u(v)$ to denote the opinion of u about v.

Now for the constraints of ϕ': For every walk w in G_ϕ starting at vertex u and ending at v of length $\ell \in [k/2, k]$ ϕ' has $D^{k-\ell}$ *copies* of the constraint $F(X'_u, X'_v)$ which imposes the conjunction of all constraints within balls of radius k of u and v. Specifically, (1) For every $u' \in B(u,k) \cap B(v,k)$ it must hold that $X'_u(u') = X'_v(u')$. (2) For $u' \in B(u,k) \cup B(v,k)$ let $O(u) = X_u(u')$ or $X_v(u')$ whichever is defined (notice by (1) that these are consistent). For every u', v' such that $u', v' \in B(u,k) \cup B(v,k)$ and $f(u', v')$ is a constraint of ϕ, it must hold that $f(O(u'), O(v')) = 1$. Notice that the many *copies* of each constraint ensure that a randomly chosen constraint of ϕ' will correspond to a walk whose length is distributed uniformly over the interval $k/2, \ldots, k$.

The above gives the complete description of the reduction essentially used in Dinur's work. We won't be able to give the proof as to why it works here. Even worse, we won't even be able to motivate the reasons behind the many delicate choices made in the reduction above. (Why are the new variables chosen as they are? Why do we create walks of so many different lengths? Why do we replicate the constraints in this way?) All we can say is that these choices are not necessarily

the first ones one may consider, but definitely make the proof of the amplification lemma very easy. The reader is encouraged to followup by reading the original paper.

7 Conclusion

We hope the reader finds the above description to be somewhat useful, and motivating when reading Dinur's new approach to consruction of PCPs. We remark that the earlier algebraic approaches, while technically much more complicated, do have some appealing high level views. The reader is pointed to the work of Ben-Sasson and this author [10] to get a sense of some of the work in the older stream.

Moving on beyond the specific proofs, and constructions used to get probabilistically checkable proofs, we hope that the notion itself is appealing to the reader. The seemingly counterintuitive properties of probabilistically checkable proofs highlight the fact the "format" in which a proof is expected is a very powerful tool to aid the person who is verifying proofs. Indeed for many computer generated proofs of mathematical theorems, this notion may ease verifiability, though in order to do so, PCPs need to get shorter than they are; and they verification scheme simpler than it is. Dinur's work helps in this setting, but much more needs to be done.

And finally, moving beyond the notion of proofs, we also hope this article reminds the reader once more of a fundamental question in logic, and computation, and indeed for all mathematics: Is P=NP? Can we really replace every mathematician by a computer? If not, would it not be nice to have a proof of this fact?

References

[1] M. Ajtai, J. Komlos, and E. Szemeredi. Deterministic simulation in logspace. In *Proceedings of the 19th Annual ACM Symposium on Theory of Computing*, pages 132–140, 1987.

[2] Sanjeev Arora, Carsten Lund, Rajeev Motwani, Madhu Sudan, and Mario Szegedy. Proof verification and the hardness of approximation problems. *Journal of the ACM*, 45(3):501–555, May 1998.

[3] Sanjeev Arora and Shmuel Safra. Probabilistic checking of proofs: A new characterization of NP. *Journal of the ACM*, 45(1):70–122, January 1998.

[4] László Babai, Lance Fortnow, Leonid A. Levin, and Mario Szegedy. Checking computations in polylogarithmic time. In *Proceedings of the 23rd ACM Symposium on the Theory of Computing*, pages 21–32. ACM, New York, 1991.

[5] László Babai, Lance Fortnow, and Carsten Lund. Non-deterministic exponential time has two-prover interactive protocols. *Computational Complexity*, 1(1):3–40, 1991.

[6] László Babai and Shlomo Moran. Arthur-Merlin games: a randomized proof system, and a hierarchy of complexity class. *Journal of Computer and System Sciences*, 36(2):254–276, April 1988.

[7] Mihir Bellare, Oded Goldreich, and Madhu Sudan. Free bits, PCP's and non-approximability — towards tight results. *SIAM Journal on Computing*, 27(3):804–915, 1998.

[8] Mihir Bellare, Shafi Goldwasser, Carsten Lund, and Alex Russell. Efficient probabilistically checkable proofs and applications to approximation. In *Proceedings of the 25th ACM Symposium on the Theory of Computing*, pages 294–304. ACM, New York, 1993.

[9] M. Ben-Or, S. Goldwasser, J. Kilian, and A. Wigderson. Multi-prover interactive proofs: How to remove intractability. In *Proceedings of the 20th Annual ACM Symposium on the Theory of Computing*, pages 113–131, 1988.

[10] Eli Ben-Sasson and Madhu Sudan. Short PCPs with poly-log rate and query complexity. In *Proceedings of the 37th Annual ACM Symposium on Theory of Computing*, pages 266–275, 2005.

[11] Aviad Cohen and Avi Wigderson. Dispersers, deterministic amplification, and weak random sources (extended abstract). In *IEEE Symposium on Foundations of Computer Science*, pages 14–19, 1989.

[12] Irit Dinur. The PCP theorem by gap amplification. Technical Report TR05-046, ECCC, 2005. Revision 1, Available from `http://eccc.uni-trier.de/eccc-reports/2005/TR05-046/`.

[13] Irit Dinur and Omer Reingold. Assignment testers: Towards a combinatorial proof of the PCP-theorem. In *Proceedings of the 45th Annual IEEE Symposium on Foundations of Computer Science*, pages 155–164, 2004.

[14] Uriel Feige, Shafi Goldwasser, Laszlo Lovasz, Shmuel Safra, and Mario Szegedy. Interactive proofs and the hardness of approximating cliques. *Journal of the ACM*, 43(2):268–292, 1996.

[15] Lance Fortnow, John Rompel, and Michael Sipser. On the power of multi-prover interactive protocols. *Theoretical Computer Science*, 134(2):545–557, 1994.

[16] M. R. Garey, D. S. Johnson, and L. J. Stockmeyer. Some simplified np-complete graph problems. *Theoretical Computer Science*, 1(3):237–267, 1976.

[17] Shafi Goldwasser, Silvio Micali, and Charles Rackoff. The knowledge complexity of interactive proof systems. *SIAM Journal on Computing*, 18(1):186–208, February 1989.

[18] Johan Håstad. Some optimal inapproximability results. *Journal of the ACM*, 48:798–859, 2001.

[19] Russell Impagliazzo and David Zuckerman. How to recycle random bits. In *IEEE Symposium on Foundations of Computer Science*, pages 248–253, 1989.

[20] Nati Linial and Avi Wigderson. Expander graphs and their applications: Lecture notes of a course given at the Hebrew University, 2003. Available from `http://www.math.ias.edu/~avi/TALKS/expander_course.ps`.

[21] Carsten Lund and Mihalis Yannakakis. On the hardness of approximating minimization problems. *Journal of the ACM*, 41(5):960–981, September 1994.

[22] Rajeev Motwani and Prabhakar Raghavan. *Randomized Algorithms*. Cambridge University Press, 1995.

[23] C.H. Papadimitriou and M. Yannakakis. Optimization, approximation, and complexity classes. *Journal of Computer and System Sciences*, 43:425–440, 1991.

[24] Ran Raz. A parallel repetition theorem. *SIAM Journal on Computing*, 27(3):763–803, 1998.

[25] Madhu Sudan. PCP and inapproximability: Survey and open problems, February 2000. Slides of Talk given at DIMACS Workshop on Approximability of NP-hard problems, Nassau Inn, Princeton, NJ. Available from `http://theory.csail.mit.edu/~madhu/slides/dimacs00.ps`.

Symmetry and Neuroscience

Martin Golubitsky[1]

1 Introduction

We discuss the question: Is symmetry a useful tool in neuroscience? Even though symmetry has been important in many aspects of physics and engineering, it may appear to be an unlikely part of the structure of the nervous system. However, there are at least three rather different areas of neuroscience where symmetry does have a role to play: animal gaits, the visual cortex, and the vestibular system; and this talk will describe how symmetries enter into the these areas. My point of view is the one discussed in *The Symmetry Perspective* [17].

As an overview Schöner, Jiang, and Kelso [32] and Collins and Stewart [10] point out that standard quadrupedal gaits (walk, trot, pace, etc.) are highly stylized symmetric motions. Collins and Stewart observe that these symmetries suggest a structure for locomotor central pattern generators. Moreover, understanding these gaits leads to interesting mathematics concerning the spatiotemporal symmetries of periodic solutions of ODEs.

Ermentrout, Cowan, and Bressloff [12, 5, 6] exploit symmetries in the connectivity of the primary visual cortex to create models for this system, and use the symmetries to explain the form that geometric visual hallucinations take. Finally, McCollum and Boyle [29] show that the neuroconnectivity between the semicircular canals in the inner ear and the ring of muscles surrounding the neck has octahedral symmetry.

In each of these examples, the symmetry group needs to be identified either through the symmetries found in system outputs (gaits, hallucinations) or in the actual neurobiology (primary visual cortex, vestibular system). In addition, in order to be useful in modeling, the spaces on which these symmetries act must be identified — and usually these spaces are understood using the phase spaces of coupled systems of ODEs [35, 23]. For example, the simplest models for quadrupedal gaits are based on the group $\mathbf{Z}_4 \times \mathbf{Z}_2$ acting on $\mathbf{R}^8$ (its right regular representation); the simplest model for orientation sensitivity in the visual cortex is given by the planar Euclidean group $\mathbf{E}(2)$ acting on itself by group multiplication; and the simplest model of the canal-neck projection of the vestibular system appears to be the octahedral group (rotational symmetries of the cube) acting on $\mathbf{R}^7$.

How important are these symmetries? That remains to be determined. But, at the very least, these are curious and interesting observations. We discuss each in turn. The presentation on gaits follows [18, 17]; the presentation on the visual cortex follows [5, 16]; and the presentation on the vestibular system follows [22].

2 Animal Locomotion

The general phenomenology of symmetric networks can be illustrated in the context of animal locomotion [7, 9, 10, 19, 20]. It has long been recognized that legged locomotion

[1]Department of Mathematics, University of Houston, Houston TX 77204-3008. E-mail: mg@uh.edu

involves a variety of standard spatio-temporal patterns, in which the legs move periodically in a particular sequence and with particular phase relationships. The case of quadrupeds is especially familiar. For example, when a horse trots, diagonally opposed legs are synchronized, but the two diagonals are half a period out of phase. When the horse walks, the legs hit the ground in the sequence left rear, left front, right rear, right front (or its left/right mirror image) at intervals of one quarter period. When a camel or giraffe paces, its left legs are synchronous, its right legs are synchronous, but left and right are half a period out of phase. More complex gaits, such as the gallop, have phase shifts that are not such simple fractions of the period, leading to a distinction between *primary* gaits with very rigid, simple phase shifts, and *secondary* gaits with more arbitrary and more flexible ones. Figure 1 shows seven common quadruped gaits. Dogs tend to walk, trot, and transverse gallop; squirrels bound; camels tend to pace and rotary gallop; and deer often pronk (all legs moving in synchrony) when startled.

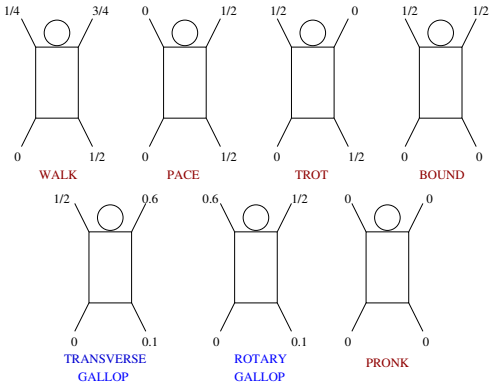


Figure 1: Seven quadrupedal gaits. Numbers indicate the percentage of the time through the gait when the associated leg first strikes the ground. Gaits begin when left hind leg strikes ground.

The symmetry approach to gaits aims to provide a rationale for these patterns, and to explain the distinction between primary and secondary gaits. It tackles these problems by seeking the schematic form of the animal's central pattern generator (CPG), see Kopell and Ermentrout [28]. The CPG is a network of neurons that is widely believed to generate nerve signals with the corresponding gait spatio-temporal rhythms. Its existence is supported by much indirect evidence, see for example Grillner and Wallén [25], but significant information on the detailed structure of the CPG is known only for a few animals, notably the lamprey, see for example Grillner *et al.* [24]. For most animals even the existence of a CPG has not been confirmed directly, though it is well established that the basic rhythms of locomotion are generated somewhere in the spinal cord, not in the brain. It therefore makes sense to try to infer qualitative information about the CPG from the gaits themselves. Such inferences

must start by making some assumptions about the nature of the CPG and how it relates to the gaits, and the consequent deductions are only as good as those assumptions.

We consider three issues: the spatiotemporal symmetries of periodic solutions to systems of ODEs with examples given by quadruped locomotion, the structure that a minimal network of coupled systems of ODEs must have in order to produce robustly periodic solutions with prescribed spatiotemporal symmetries, and a prediction made by this minimal model.

Coupled Cell Networks

For the purpose of this talk a *coupled cell system* is a collection of identical systems of ODE, or *cells*, that are identically coupled. The *network* is a graph whose nodes are the cells and whose arrows indicate which cells are coupled to which. The beginnings of a general theory for the dynamics of coupled cell networks has been developed in [35, 23, 18].

The simplest coupled cell network is the two-cell one in Figure 2. We associate with this network a class of differential equations, which we call *admissible*. For this network the admissible differential equations are those of the form

$$\begin{aligned} \dot{x}_1 &= g(x_1, x_2) \\ \dot{x}_2 &= g(x_2, x_1) \end{aligned} \tag{2.1}$$

where $x_1, x_2 \in \mathbf{R}^k$ are the state variables of the individual cells. Observe that a single function $g : \mathbf{R}^k \times \mathbf{R}^k \to \mathbf{R}^k$ defines the system.

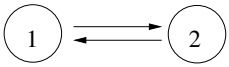


Figure 2: Schematic of a two identical cell identical coupling network.

One consequence of the $\mathbf{Z}_2$ symmetry of this system is the existence of solutions in which $x_1(t) = x_2(t)$ for all t. This follows because the diagonal subspace $\{x : x_1 = x_2\}$ is invariant under the flow of the differential equation for all g. For all such solutions, the two cells behave synchronously. In particular, there is a nonempty open set of functions g for which these systems have synchronous periodic solutions.

Another consequence of symmetry is that there exists a non-empty open set of functions g for which there is a periodic solution, with period T, such that $x_2(t) = x_1(t + T/2)$ for all t, see [21, 17]. That is, the two cells have the same periodic dynamics except for a relative phase shift of half a period. The existence of these two types of periodic solutions generalizes to the class of admissible vector fields for symmetric networks as follows.

Spatiotemporal Symmetries of Periodic Solutions

A symmetry of an ordinary differential equation (ODE) is a transformation that sends solutions to solutions. More specifically, let $\gamma : \mathbf{R}^n \to \mathbf{R}^n$ be a linear map. A system of

differential equations

$$\dot{x} = f(x) \tag{2.2}$$

(where $x \in \mathbf{R}^n$ and $f : \mathbf{R}^n \to \mathbf{R}^n$ is smooth) has *symmetry* γ if $\gamma x(t)$ is a solution to (2.2) whenever $x(t)$ is a solution. It is straightforward to verify that γ is a symmetry if and only if f satisfies the *equivariance condition*

$$f(\gamma x) = \gamma f(x) \tag{2.3}$$

Suppose that the system (2.2) has a finite symmetry group Γ. We first note that symmetry forces the existence of many flow invariant subspaces. Suppose that $\Sigma \subset \Gamma$ is a subgroup. Then

$$\mathrm{Fix}(\Sigma) = \{x \in \mathbf{R}^n : \sigma x = x \ \ \forall \sigma \in \Sigma\}$$

is a flow-invariant subspace. (Proof: $\sigma f(x) = f(\sigma x) = f(x)$ for each $x \in \mathrm{Fix}(\Sigma)$. Hence $f : \mathrm{Fix}(\Sigma) \to \mathrm{Fix}(\Sigma)$.) In case Γ is the symmetry group of a network (that is, Γ is a permutation group of the cells), the fixed-point subspaces are generalized diagonals and flow-invariance implies synchrony.

Second, phase-locking is also a natural consequence of symmetry. Suppose that $x(t)$ is a T-periodic solution to (2.2) and that γ is a symmetry. Then either $\gamma x(t)$ is a different periodic trajectory from $x(t)$, or it is the same trajectory. In the latter case, the only difference is a time-translation. That is, $\gamma x(0) = x(\theta)$ for some θ, and uniqueness of solutions implies that $\gamma x(t) = x(t + \theta)$ for all t. Define

$$\begin{array}{rcll} H & = & \{\gamma \in \Gamma : \gamma\{x(t)\} = \{x(t)\}\} & \text{spatiotemporal symmetries} \\ K & = & \{\gamma \in \Gamma : \gamma x(t) = x(t) \quad \forall t\} & \text{spatial symmetries} \end{array}$$

Note that since fixed-point subspaces are flow-invariant, K is an isotropy subgroup of the action of Γ on $\mathbf{R}^n$. In addition, for each $h \in H$, there is a phase shift $\theta(h) \in \mathbf{S}^1$ such that $hx(t) = x(t + \theta(h))$. Moreover, $\theta : H \to \mathbf{S}^1$ is a group homomorphism with kernel K. It follows that H/K is isomorphic to a finite subgroup of $\mathbf{S}^1$ and hence is cyclic.

Periodic solutions with spatiotemporal symmetries are classified as follows.

Theorem 2.1 (H/K Theorem [7, 17]) *Let Γ be a prmutation group which is the symmetry group of a coupled cell network in which all cells are coupled and the internal dynamics of each cell is at least two-dimensional. Let $K \subset H \subset \Gamma$ be a pair of subgroups. Then there exist periodic solutions to some coupled cell system with spatiotemporal symmetries H and spatial symmetries K if and only if H/K is cyclic and K is an isotropy subgroup.*

Four Cells Do Not Suffice

The simplest model of a quadruped locomotor CPG has four identical cells, where it is presumed that the output signal from each cell is sent to one leg. See Figure 3. We ask whether it is possible to couple these four cells in such a way that network systems can naturally produce rhythms associated with the three gaits walk, trot, and pace, and show that it is not [7].

To justify this negative statement we discuss three points:

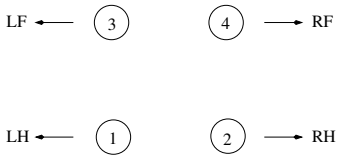


Figure 3: Signal from cell 1 is sent to left hind (LH) leg, etc.

(a) Gaits rhythms are described by spatiotemporal symmetries.

(b) The symmetry groups of trot and pace cannot be conjugate.

(c) The symmetry group of trot and pace are always conjugate in any four-cell network that also produces a walk.

(a) Collins and Stewart [10] observed that standard quadruped gaits are distinguished by spatiotemporal symmetries, where the space symmetries are leg permutations. The generators for the symmetry groups of trot, pace, and walk are listed in Table 1. In our models we assume that gait rhythms are exact and robust. We also assume that the only robust phase shifts of periodic solutions that are given in these models are those that are described by symmetry.

Gait	Generators of spatio-temporal symmetries			Solution form
Trot	$((1\ 2)(3\ 4), \frac{1}{2})$	and	$((1\ 3)(2\ 4), \frac{1}{2})$	$(x(t), x(t+\frac{1}{2}), x(t+\frac{1}{2}), x(t))$
Pace	$((1\ 2)(3\ 4), \frac{1}{2})$	and	$((1\ 3)(2\ 4), 0)$	$(x(t), x(t+\frac{1}{2}), x(t), x(t+\frac{1}{2}))$
Walk	$((1\ 3\ 2\ 4), \frac{1}{4})$			$(x(t), x(t+\frac{1}{2}), x(t+\frac{1}{4}), x(t+\frac{3}{4}))$

Table 1: Legs are numbered by the associated cells in Figure 3. The permutation (1 2)(3 4) swaps left and right legs; the permutation (1 3)(2 4) swaps front and back legs; fractions indicate phase shift as a fraction of a gait period.

(b) Experiments on dogs imply that trot and pace are not gaits that can be modeled by conjugate solutions. Note that in a system of differential equations conjugate solutions differ only by initial conditions and have the same stability. Blaszczyk and Dobrzecka [2] indicate that the stability of pace and trot are not the same. In their experiment, a dog's legs are restrained so that they can use a pace at intermediate speeds, but not a trot, which is the dog's preferred gait. Different dogs are placed in this device for two to six months. In post-restraint trials dogs that were in the shorter restraint period switched back to a trot quickly with only occasional use of a pace. Occurrence of the pace was more frequent in the animals that were restrained for a longer period, but the use of pace decreased with every post-restraint experimental trial.

(c) It follows from (a) that if a four-cell network is coupled so that periodic solutions with the rhythm of a walk occur naturally, then the permutation (1 3 2 4) must be a

network symmetry. Suppose that the system also produces a pace solution. As indicated in Figure 4, cells 1 and 3 and cells 2 and 4 must be synchronous. As illustrated in that figure, applying the walk symmetry to that solution produces a solution in which cells 1 and 4 and cells 2 and 3 are synchronous — a pace. It follows that trot and pace solutions are conjugate in any four-cell network that can produce a robust walk.

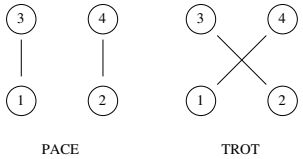


Figure 4: Lines between cells indicate synchrony; no lines indicate half-period phase shifts.

The Eight-Cell Network

Golubitsky *et al.* [7, 20] make six assumptions, and deduce that for quadrupeds the only possible symmetry class of CPG networks is the 8-cell network shown in Figure 5. The details of the deduction are unimportant here, but they are explicit in the original paper.

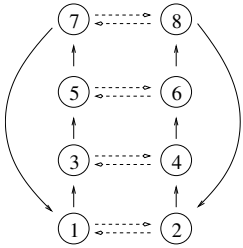


Figure 5: Eight-cell network for quadrupeds. Dashed lines indicate contralateral coupling; single lines indicate ipsilateral coupling.

This network has eight symmetries—permutations of the legs (more precisely, the leg labels) that preserve the edges. There are two types of symmetry: contralateral symmetry κ, which interchanges cells on the left with cells on the right, and ipsilateral symmetry ω, which cyclically and simultaneously permutes cells on both left and right. Thus the symmetry group of the eight-cell quadruped CPG is $\Gamma = \mathbf{Z}_2\langle\kappa\rangle \times \mathbf{Z}_4\langle\omega\rangle$.

The H/K Theorem provides a classification of the possible spatio-temporal symmetries. Primary states are characterized by all eight cells having the same waveform modulo phase shift (that is, $H = \Gamma$) whereas secondary gaits involve more than one waveform (that is, $H \subsetneq \Gamma$). It is straightforward to calculate the six subgroups $K \subset H$ for which H/K is

		walk	jump	trot	pace	bound	pronk
LF	RF	$\frac{3}{4}$ $\frac{1}{4}$	$\frac{1}{2}$ $\frac{1}{2}$	$\frac{1}{2}$ 0	0 $\frac{1}{2}$	$\frac{1}{2}$ $\frac{1}{2}$	0 0
LH	RH	$\frac{1}{2}$ 0	$\frac{3}{4}$ $\frac{3}{4}$	0 $\frac{1}{2}$	0 $\frac{1}{2}$	0 0	0 0
LF	RF	$\frac{1}{4}$ $\frac{3}{4}$	0 0	$\frac{1}{2}$ 0	0 $\frac{1}{2}$	$\frac{1}{2}$ $\frac{1}{2}$	0 0
LH	RH	0 $\frac{1}{2}$	$\frac{1}{4}$ $\frac{1}{4}$	0 $\frac{1}{2}$	0 $\frac{1}{2}$	0 0	0 0
Subgroup K		$\mathbf{Z}_2(\kappa\omega^2)$	$\mathbf{Z}_2(\kappa)$	$\mathbf{Z}_4(\kappa\omega)$	$\mathbf{Z}_4(\omega)$	$\mathbf{D}_2(\kappa,\omega^2)$	$\mathbf{Z}_2\times\mathbf{Z}_4$

Table 2: Phase shifts for primary gaits in the eight-cell network.

cyclic and determine the primary patterns for the 8-cell network: see Table 2. There is an analogous (but more complicated) classification of secondary gaits.

The fact that $H = \Gamma$ implies that the signals $x_i(t)$ and $x_j(t)$ must be the same up to a well defined phase shift. For example, suppose that K is generated by κ and ω^2. Since κ is a K symmetry the outputs from κ related cells must be identical; that is, $x_1(t) = x_2(t)$, etc. Since ω^2 is a K symmetry, $x_1(t) = x_5(t)$, etc. Since ω is an H symmetry that is not in K, it corresponds to a half period phase shift and $x_1(t) = x_3(t + \frac{1}{2})$. For such a periodic solution this model CPG sends synchronous signals to the hind legs, synchronous signals to the fore legs, and the two sets of signals are a half period out of phase. This rhythm corresponds to a bound. The other identifications with gait rhythms are found similarly.

A Prediction: The Jump

The patterns listed in the table correspond to standard primary quadruped gaits, with one exception: the gait we have labelled 'jump'. After performing the above analysis, the jump gait was observed at the Houston Livestock Show and Rodeo. Figure 6 shows four video frames of a bucking bronco, taken at equal intervals of time. The interval between the footfalls is very close to 1/4 of the period of this rhythmic motion.

Indeed, approximately 200 frames of the rodeo video are coded in Figure 7. Dark regions begin when the right hind leg is firmly on the ground and light regions begin when the right fore leg is firmly on the ground. This figure indicates that the average time elapsed from right hind to right fore leg ground strikes is approximately three times the average time elapsed from right fore to right hind leg ground strikes. The *primitive ricocheting jump* of a Norway rat and an Asia Minor gerbil also has the same pattern of phases as the jump gait, Gambaryan [14].

3 The Primary Visual Cortex

The orientation sensitivity of neurons in the primary visual cortex appears to encode the Euclidean group, acting on itself by group multiplication, as a group of symmetries of the cortex, and these symmetries appear to characterize the kinds of geometric patterns described by individuals undergoing drug-induced visual hallucinations. In earlier work, Bard

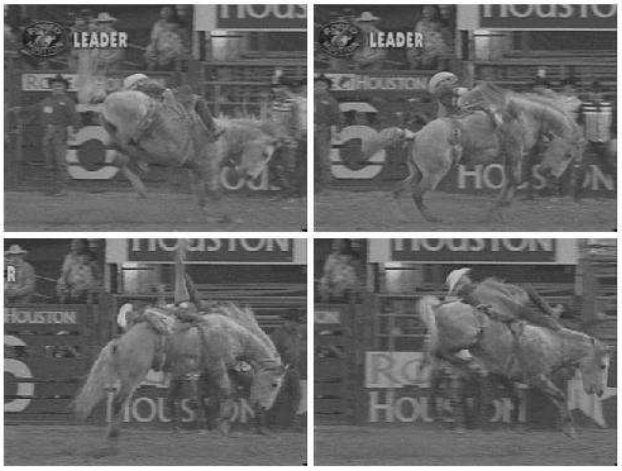


Figure 6: Quarter cycles of bareback bronc jump at Houston Livestock Show and Rodeo. (UL) fore legs hit ground; (UR) hind legs hit ground; (LL) and (LR) all legs in air.

Ermentrout and Jack Cowan [12] proposed explaining the geometric forms of hallucinations by applying ideas from equivariant bifurcation theory to continuum models of the visual cortex. In later work Cowan, along with Paul Bresloff, Peter Thomas, and Matt Wiener [5, 6], proposed using the orientation sensitivity of neurons in the primary visual cortex to refine the symmetry arguments and to obtain results that better coordinated the mathematically generated patterns with the drug induced images.

A Short Review of Geometric Hallucinations

In the 1930's Klüver classified geometric visual hallucinations into four groups of *form constants* (see [27, p. 66]): honeycombs, cobwebs, tunnels, and spirals. Klüver states on p. 71 "We wish to stress merely one point, namely, that under diverse conditions the visual system responds in terms of a limited number of form constants." Examples of the four form constants are given in Figure 8.

Ermentrout and Cowan [12] pioneered an approach to the mathematical study of geometric patterns produced in drug induced hallucinations. They assumed that the drug uniformly stimulates an inactive cortex and produces, by spontaneous symmetry-breaking, a patterned activity state. The mind then interprets the pattern as a visual image — namely the visual image that would produce the same pattern of activity on the primary visual cortex V1. The Ermentrout-Cowan analysis assumes that a differential equation

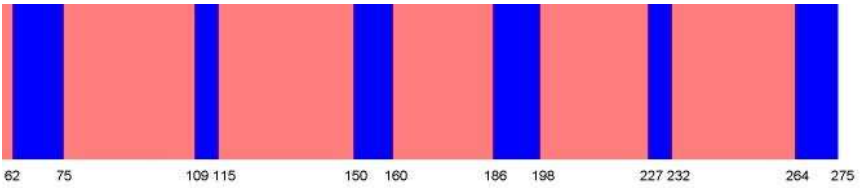


Figure 7: Average right hind to right fore = 31.2 frames (light region); average right fore to right hind = 11.4 frames (dark region); $\frac{31.2}{11.4} = 2.74$.

governs the symmetry-breaking transition from an inactive to an active cortex and then studies abstractly the transition using standard pattern formation arguments developed for reaction-diffusion equations [21, 17]. Their cortical patterns are obtained by thresholding (points where the solution is greater than some threshold are colored black, whereas all other points are colored white). These cortical patterns are then transformed to retinal patterns using the inverse of the retino-cortical map described below (see (3.3)), and these retinal patterns are similar to some of the geometric patterns of visual hallucinations, namely, funnels and spirals.

Orientation Sensitvity of Neurons in the Visual Cortex

It is now well established that neurons in V1 are sensitive to orientations in the visual field. See [26, 15, 1, 3, 5] for more discussion. It is mathematically reasonable to assign an orientation preference to each neuron in V1. Hubel and Wiesel [26] introduced the notion of a *hypercolumn* — a region in V1 containing for each orientation at a single point in the visual field (a mathematical idealization) a neuron sensitive to that orientation.

More recently, Bressloff *et al.* [5] studied the geometric patterns of drug induced hallucinations by including orientation sensitivity. As before, the drug stimulation is assumed to induce spontaneous symmetry-breaking, and the analysis is local in the sense of bifurcation theory. There is one major difference between the approaches in [5] and [12]. Ignoring lateral boundaries Ermentrout and Cowan [12] idealize the cortex as a plane, whereas Bressloff *et al.* [5] take into account the orientation tuning of cortical neurons and idealize the cortex as $\mathbf{R}^2 \times \mathbf{S}^1$. This approach leads to a method for recovering thin line hallucinations such as cobwebs and honeycombs, in addition to the threshold patterns found in the Ermentrout-Cowan theory. See Figure 9.

There are two types of connections between neurons in V1: local and lateral. Experimental evidence suggests that neurons within a hypercolumn are all-to-all connected, whereas neurons in different hypercolumns are connected in a very structured way. This structured lateral coupling is called *anisotropic*, and it is the bifurcation theory associated with anisotropic coupling that is studied in Bressloff *et al.* [5, 4].

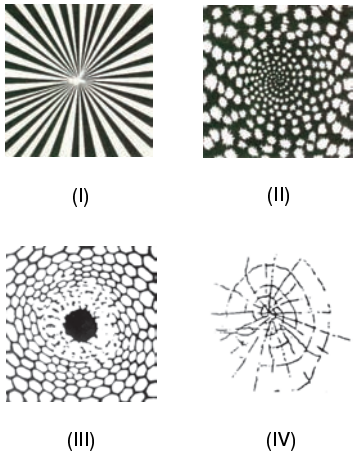


Figure 8: Hallucinatory form constants from [6]. (I) funnel and (II) spiral images seen following ingestion of LSD [redrawn from [33], (III) honeycomb generated by marihuana [redrawn from [8]], (IV) cobweb petroglyph [Redrawn from [30]].

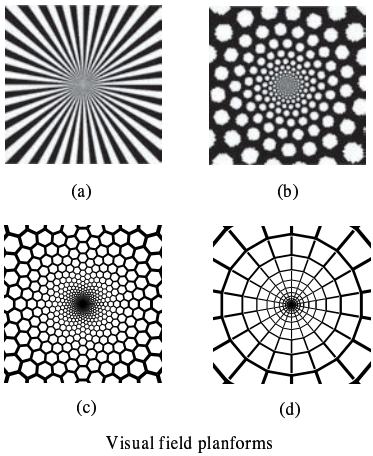


Figure 9: Hallucinatory form constants generated by symmetry-breaking bifurcations on cortex using the shift-twist representation of the Euclidean group, and viewed in retinal coordinates. From [6]. Note the similarities with Figure 8.

The Continuum Models

The Ermentrout and Cowan [12] model of V1 consists of neurons located at each point $\mathbf{x}$ in $\mathbf{R}^2$. Their model equations, variants of the Wilson-Cowan equations [37], are written in terms of a real-valued *activity variable* $a(\mathbf{x})$, where a represents, say, the voltage potential of the neuron at location $\mathbf{x}$.

Bressloff *et al.* [5] incorporate the Hubel-Weisel hypercolumns [26] into their model of V1 by assuming that there is a hypercolumn centered at each location $\mathbf{x}$. Here a *hypercolumn* denotes a region of cortex that contains neurons sensitive to orientation ϕ for each direction ϕ. Their models, also adaptations of the Wilson-Cowan equations [37], are written in terms of a real-valued *activity variable* $a(\mathbf{x}, \phi)$ where a represents, say, the voltage potential of the neuron tuned to orientation ϕ in the hypercolumn centered at location $\mathbf{x}$. Note that angles ϕ and $\phi + \pi$ give the same orientation; so $a(\mathbf{x}, \phi + \pi) = a(\mathbf{x}, \phi)$.

The cortical planform associated to $a(\mathbf{x}, \phi)$ is obtained in a way different from the Ermentrout-Cowan approach. For each fixed $\mathbf{x} \in \mathbf{R}^2$, $a(\mathbf{x}, \cdot)$ is a function on the circle. The planform associated to a is obtained through a *winner-take-all* strategy. The neuron that is most active in its hypercolumn is presumed to suppress the activity of other neurons within that hypercolumn. The winner-take-all strategy chooses, for each $\mathbf{x}$, the directions ϕ that maximize $a(\mathbf{x}, \cdot)$, and results in a field of directions. The two approaches to creating planforms can be combined by assigning directions only to those locations $\mathbf{x}$ where the associated maximum of $a(\mathbf{x}, \cdot)$ is larger than a given threshold.

A possible justification for the continuum model that idealizes a hypercolumn at each cortex location is that each location is in fact surrounded by neurons sensitive to all of the possible orientations. This fact suggests that the signal read from the primary visual cortex V1 need not be limited to one orientation from each 'physical' hypercolumn. In V1 there is a grid of physical hypercolumns that is approximately 36×36 in extent. (See [4] and references therein.) It is reasonable to suppose that other layers of the visual cortex receive much more information than a 36×36 matrix of orientation values.

Euclidean Symmetry

The Euclidean group $\mathbf{E}(2)$ is crucial to the analyses in both [12] and [5] — but the way that group acts is different. In Ermentrout-Cowan the Euclidean group acts on the plane by its standard action, whereas in Bressloff *et al.* the Euclidean group acts on $\mathbf{R}^2 \times \mathbf{S}^1$ by the so-called shift-twist representation, as we now explain.

Bressloff *et al.* [5] argue, based on experiments by Blasdel [1] and Eysel [13], that the lateral connections between neurons in neighboring hypercolumns are *anisotropic*. That anisotropy states that the *strength* of the connections between neurons in two neighboring hypercolumns depends on the orientation tuning of both neurons and on the relative locations of the two hypercolumns. Moreover, this anisotropy is idealized to the one illustrated in Figure 10 where only neurons with the same orientation selectivity are connected and then only neurons that are oriented along the direction of their cells preference are connected. These conclusions are based on work of Gilbert [15] and Bosking *et al.* [3]. In

particular, the symmetries of V1 model equations are those that are consistent with the idealized structure shown in Figure 10.

The Euclidean group $\mathbf{E}(2)$ is generated by translations, rotations, and a reflection. The action of $\mathbf{E}(2)$ on $\mathbf{R}^2 \times \mathbf{S}^1$ that preserves the structure of lateral connections illustrated in Figure 10 is the *shift-twist* action. This action is given by:

$$\begin{array}{rcl} \mathcal{T}_{\mathbf{y}}(\mathbf{x}, \phi) & \equiv & (\mathbf{x} + \mathbf{y}, \phi) \\ \mathcal{R}_{\theta}(\mathbf{x}, \phi) & \equiv & (R_{\theta}\mathbf{x}, \phi + \theta) \\ \mathcal{M}_{\kappa}(\mathbf{x}, \phi) & \equiv & (\kappa\mathbf{x}, -\phi), \end{array} \tag{3.1}$$

where $(\mathbf{x}, \phi) \in \mathbf{R}^2 \times \mathbf{S}^1$, $\mathbf{y} \in \mathbf{R}^2$, κ is the reflection $(x_1, x_2) \mapsto (x_1, -x_2)$, and $R_{\theta} \in \mathbf{SO}(2)$ is rotation of the plane counterclockwise through angle θ.

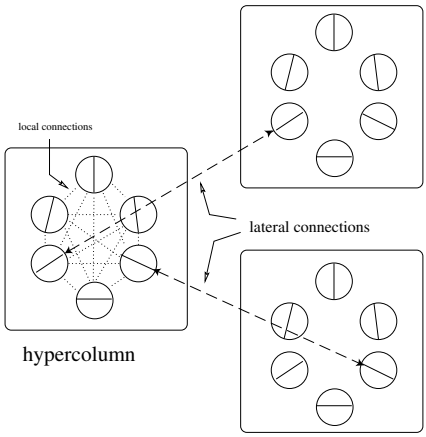


Figure 10: Illustration of isotropic local and anisotropic lateral connection patterns.

Work on optical imaging has made it possible to see how the orientation preference of cells are actually distributed in V1 [1], and a variety of stains and labels have made it possible to see how they are interconnected [13, 3]. Figure 11 shows that the distribution of orientation preferences in the Macaque. In particular, approximately every millimeter there is an *iso-orientation patch* of a given preference.

Recent optical imaging experiments combined with anatomical tracer injections suggest that there is a spatial anisotropy in the distribution of patchy horizontal connections, as illustrated in Figure 12. It will be seen from the right panel that the anisotropy is particularly pronounced in the tree shrew. The major axis of the horizontal connections tends to run parallel to the visuotopic axis of the connected cells' common orientation preference. There is also a clear anisotropy in the patchy connections of Macaque, as seen in the left panel.

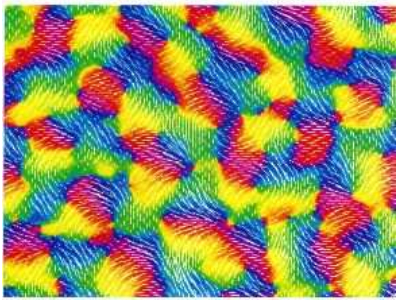


Figure 11: Distribution of orientation preferences in Macaque V1 obtained via optical imaging and using color to indicate iso-orientation patches. Redrawn from [1].

Symmetry-Breaking Bifurcations on Lattices

Spontaneous symmetry-breaking in the presence of a noncompact group such as the Euclidean group is far from completely understood. The standard approach is to reduce the technical difficulties by looking only for solutions that are spatially doubly periodic with respect to some planar lattice (see [17]); this is the approach taken in [12, 5] and in this study. This approach is justified by the remarkable similarities between the geometric patterns obtained mathematically in [12, 5] and the hallucinatory images reported in the scientific literature [5, 6]. See Figures 8 and 9.

The first step in such an analysis is to choose a lattice type; say a square or hexagonal lattice. The second step is to decide on the size of the lattice. Euclidean symmetry guarantees that at bifurcation, critical eigenfunctions will have *plane wave* factors $e^{2\pi i \mathbf{k}\cdot\mathbf{x}}$ for some critical dual wave vector $\mathbf{k}$. See [4] or [17, Chapter 5]. Typically, the lattice size is chosen so that the critical wave vectors will be vectors of shortest length in the dual lattice; that is, the lattice has the smallest possible size that can support doubly periodic solutions.

By restricting the bifurcation problem to a lattice, the group of symmetries is transformed to a compact group. First, translations in $\mathbf{E}(2)$ act modulo the spatial period (which we can take to be 1 on the square lattice) and thus act as a 2-torus $\mathbf{T}^2$. Second, only those rotations and reflections in $\mathbf{E}(2)$ that preserve the lattice (namely, the holohedry $\mathbf{D}_4$ for the square lattice) are symmetries of the lattice restricted problem. Thus, the symmetry group of the square lattice problem is $\Gamma = \mathbf{D}_4 \dot{+} \mathbf{T}^2$. Recall that at bifurcation Γ acts on the kernel of the linearization, and a subgroup of Γ is *axial* if its fixed-point subspace in that kernel is one-dimensional. Solutions are guaranteed by the Equivariant Branching Lemma (see [21, 17]) which states: generically there are branches of equilibria to the nonlinear differential equation for every axial subgroup of Γ. The nonlinear analysis in [4, 12] proceeds in this fashion.

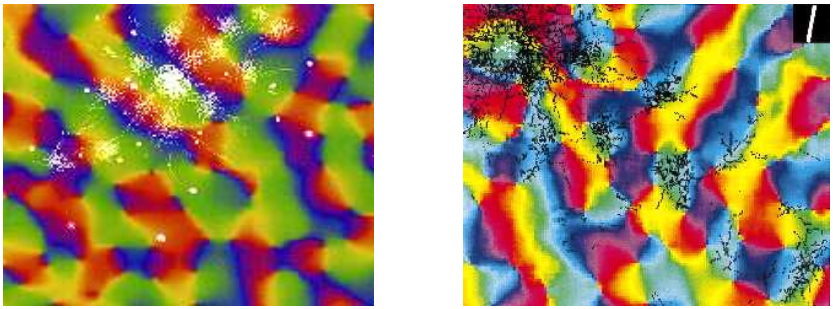


Figure 12: Lateral Connections made by a cells in Macaque (Left panel) and Tree Shrew (Right panel) V1. A radioactive tracer is used to show the locations of all terminating axons from cells in a central injection site, superimposed on an orientation map obtained by optical imaging. Redrawn from [34] and [3].

Retinal Images

Finally, we discuss the geometric form of the cortical planforms in the visual field, that is, we try to picture the corresponding visual hallucinations. It is known that the density of neurons in the visual cortex is uniform, whereas the density of neurons in the retina fall offs from the fovea at a rate of $1/r^2$. Schwartz [31] observed that there is a unique conformal map taking a disk with $1/r^2$ density to a rectangle with uniform density, namely, the complex logarithm. This is also called the *retino-cortical* map. It is thought that using the inverse of the retino-cortical map, the complex exponential, to push forward the activity pattern from V1 to the retina is a reasonable way to form the hallucination image — and this is the approach used in Ermentrout and Cowan [12] and in Bressloff *et al.* [5, 6]. Specifically, the transformation from polar coordinates (r, θ) on the retina to cortical coordinates (x, y) is given in Cowan [11] to be:

$$x = \frac{1}{\varepsilon} \ln\left(\frac{1}{\omega} r\right) \quad \text{and} \quad y = \frac{1}{\varepsilon}\theta \tag{3.2}$$

where ω and ε are constants. See Bressloff *et al.* [6] for a discussion of the values of these constants. The inverse of the retino-cortical map (3.2) is

$$r = \omega \exp(\varepsilon x) \quad \text{and} \quad \theta = \varepsilon y \tag{3.3}$$

In the retinal images, $\omega = 30/e^{2\pi}$ and $\varepsilon = 2\pi/n_h = \pi/18$, where n_h is the number of hypercolumn widths in the cortex, which is taken to be 36.

4 The Vestibular System

In this section we discuss results of McCollum and Boyd [29] as described in Golubitsky, Shiau, and Stewart [22] on the vestibular system, which is a system of tubes with sensors that sense balance and motion. There are two main components: the otolith organs, which sense linear acceleration of the head (translation), and the semicircular canals, which sense angular acceleration of the head (rotation). Each ear contains three semicircular canals arranged in three approximately mutually orthogonal planes. See Figure 13.

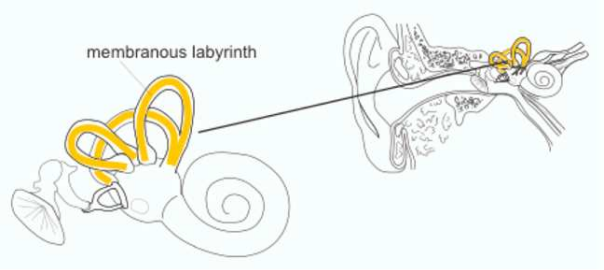


Figure 13: Structure and location of the semicircular canals (right ear). From Vilis [36].

We focus on one aspect of the vestibular system explored by McCollum and Boyd [29]: the network of neurons that conveys signals from the canals to eight principal muscle groups that control the position of the neck, known as the 'canal-neck projection'. The precise structure of the canal-neck projection network appears not to be known, but there is sufficient information to determine its connectivity and hence its symmetries, in an idealized form. Indeed, McCollum and Boyle [29] show that the canal-neck projection has the symmetries of the octahedral group $\mathbb{O}$ — the symmetry group of a cube. This group contains 48 elements, of which 24 are rotations (in the usual action by rigid motions in $\mathbf{R}^3$) and the other 24 reverse orientation.

We rederive the symmetries in the disynaptic canal-neck projection discussed by McCollum and Boyle [29]. In this aspect of the vestibular system there are six semicircular canals (three in each ear) that are connected to eight muscle groups in the neck.

Polarity Pairs of Canals

The three semicircular canals located in each ear are called *horizontal* h, *anterior* a, and *posterior* p. We denote the six canals by lh, la, lp, rh, ra, rp, where l stands for *left* and r for *right*. Canal hairs are arranged so that fluid flow in one direction in the canal stimulates an excitatory signal and fluid flow in the opposite direction stimulates an inhibitory signal. Moreover, the semicircular canals are paired (lh-rh, la-rp, lp-ra) so that when one member of a pair is transmitting an excitatory signal, then the other member of that pair is transmitting an inhibitory one. These pairs are called *polarity pairs*.

The spatial arrangement of the canals is as follows. There are three (approximately) mutually orthogonal planes. One of these planes is horizontal. The other two are vertical, at an angle of 45° to the plane of left-right symmetry of the head. Each polarity pair consists of two canals that are parallel to one of these planes: one canal in the left ear, one in the right. These two canals are oriented in opposite directions in that plane and detect rotations (actually angular accelerations) of the head about an axis perpendicular to that plane. One member of the polarity pair detects acceleration in one orientation (clockwise or counterclockwise) and the other member detects the opposite orientation.

Connections between Canals and Muscles

Each of the six canals can transmit signals to each of the eight muscle groups. The muscles form four pairs, and if a canal is activated by the motion of the head then it sends an inhibitory signal to one member of each pair and an excitatory signal to the other member. Physiological investigations suggest that each muscle group is excited by a set of three mutually orthogonal canals (that is, one from each polarity pair) and inhibited by the complementary set of canals (the other members of the polarity pairs). We describe the details of this arrangement.

Following McCollum and Boyle, the list of signals transmitted to a given muscle group can be depicted as an 'asterisk', as in Figure 14. Continuous lines represent excitatory signals and dashed lines represent inhibitory signals. Each asterisk has three solid lines (excitatory) and three dotted lines (inhibitory) and diametrically opposite lines have opposite polarity. There are eight possible arrangements of this type. Because the asterisks are drawn in two-dimensional projection, in a conventional orientation with lh between la and lp, there appear to be two kinds of asterisks: two alternating (with excitation and inhibition alternating) and six non-alternating (with three contiguous excitatory canals). We will shortly see that under a suitable action of the octahedral group, all eight asterisks are equivalent.

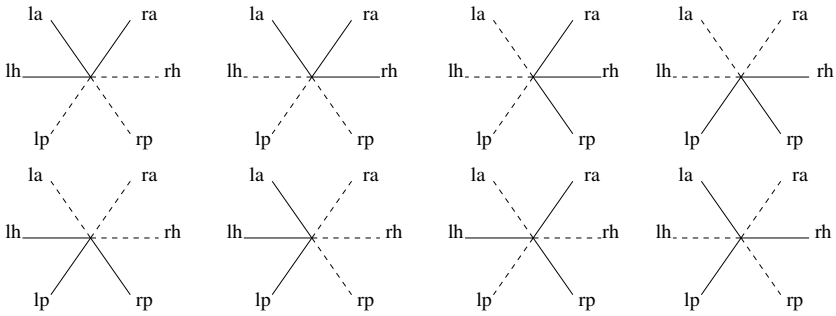


Figure 14: Eight asterisk patterns from six semicircular canals. Continuous lines represent excitatory signals and dashed lines represent inhibitory signals.

The eight neck muscles are shown in Figure 15 and consist of two flexors in the front, two extensors in the back, and four side (shoulder) muscles. The side muscles are alternating or directed. McCollum and Boyle [29] discuss the innervation patterns between canal neurons and muscle motoneurons—how the six canal neurons connect to the eight muscle motoneurons, and whether the connection occurs via an excitatory synapse or an inhibitory one. The pattern of connections to each muscle is given in Figure 16. It is important to understand that in Figure 16 an asterisk represents a list of the connections from canals to muscle groups, and type of signal that is transmitted along each connection.

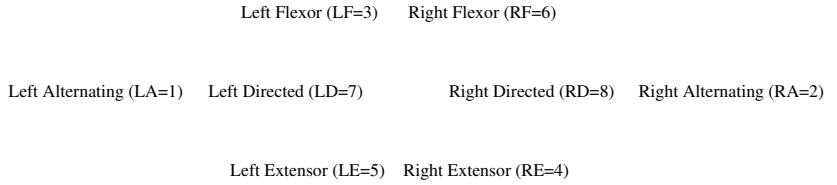


Figure 15: Location and numbering of eight muscle groups.

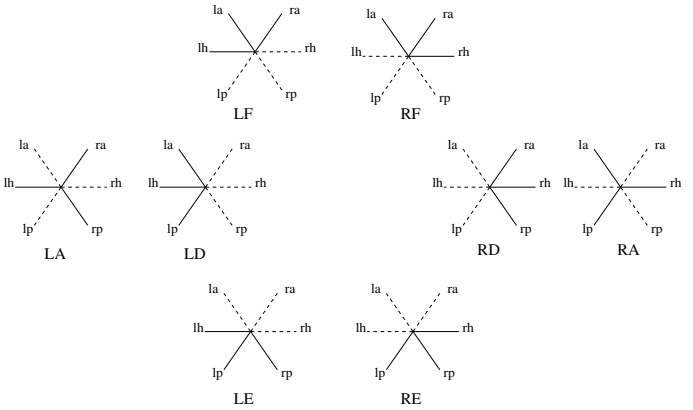


Figure 16: Innervation patterns corresponding to eight muscle groups.

Observe that the muscle groups also partition into four polarity pairs:

$$\{LA, RA\} \quad \{LF, RE\} \quad \{LE, RF\} \quad \{LD, RD\}.$$

If one muscle in a polarity pair can and does receive an excitatory signal from a canal, then the other muscle in that polarity pair can and will receive an inhibitory signal from that canal.

We illustrate the same information in another way. McCollum and Boyle [29] consider only the *disynaptic pathway* from the six vestibular nerve afferents ('canal nerves') to the eight neck motoneurons (by way of the corresponding vestibulospinal neurons). They remark that almost always 'the motoneurons of each tested muscle responded to stimulation of all six canal nerves'. The responses were classified as either excitatory or inhibitory, as indicated by solid or dotted lines for the relevant arm of the asterisk. This description makes it clear that their Figure 4 is a diagram determining these *connections*.

Octahedral symmetry of canals and muscles

McCollum and Boyle [29] show that the symmetry group of the canal-neck projection is the 48-element octahedral group $\mathbb{O}$, consisting of the 24 rotations and the 24 reflections of the cube.

It is convenient to employ a geometric image in which the canals are identified with the six faces of a cube, and the muscles with the eight vertices. The group $\mathbb{O}$ acts naturally on this picture by rigid motions of the cube in $\mathbf{R}^3$. The canals are identified with faces so that the canal polarity pairs are identified with pairs of opposite faces. Up to symmetry there is precisely one way to make this identification. Each vertex of the cube is in the intersection of exactly three faces. We identify a vertex with the asterisk whose excitatory signals correspond to the three adjoining faces. For example there is a unique vertex that is in the intersection of the three left canals (see Figure 17). We identify this vertex with the left direct muscle LD in Figure 16, since that muscle responds to excitatory signals from the three left canals.

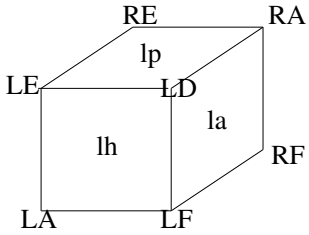


Figure 17: Identification of polarity pairs and muscle groups to the cube.

Figure 17 can also be used to construct a schematic for the connections from canals to muscle groups. Each face of the cube (a canal) is connected to the four muscle groups corresponding to a vertex on that face by a connection that innervates that muscle group when an excitatory signal is sent from the canal. The four muscle groups corresponding to vertices on the opposite face are innervated by inhibitory signals sent from that canal. It follows that any symmetry of the cube permutes canals with canals and muscle groups with muscle groups in such a way that it preserves connections and their types. Thus, in this sense, the octahedral group is the group of symmetries of the canal-neck projection.

A Minimal Phase Space for Dynamics

As we saw with the examples of gaits and the visual cortex, any mathematical analysis based on symmetry can proceed only after the symmetry group and its action on phase space have been identified. The McCollum-Boyle construction determines the symmetry group of the canal-neck projection. As discussed in [22] there is a guess as to the simplest reasonable phase space for the dynamics of this projection and hence for the representation of $\mathbb{O}$ on that phase space.

The simplest possible phase space is 14-dimensional consisting of one dimension for each canal and one dimension for each muscle group. However, the states of the canal-neck system lie in a subspace of $\mathbf{R}^{14} = \mathbf{R}^6 \times \mathbf{R}^8$. In particular, it is reasonable to identify the ith canal variable ω_i with the angular acceleration measured by that canal. Suppose that the first and second canals form a polarity pair. Since polarity pairs of canals measure opposite accelerations, we have that $\omega_2 = -\omega_1$. Similarly, it is reasonable (because of symmetry) that states of polarity pairs of muscle groups have values that are the negative of each other. Thus, the minimal state space of the canal-neck projection is $\mathbf{R}^3 \times \mathbf{R}^4 \equiv \mathbf{R}^7$.

Note that the octahedral group has the form $\mathbb{O} = \mathbf{S}_4 \oplus \mathbf{Z}_2(-I)$, where the permutation group $\mathbf{S}_4$ corresponds to the rotational symmetries of the cube. The discussion above suggests that $-I$ should permute polarity pairs of canals and muscles and multiply the result by -1. It follows that the fixed-point subspace $\mathrm{Fix}(-I)$ of the action of $-I$ on $\mathbf{R}^{14}$ is the 7-dimensional state space we have just described. Since $\mathbf{S}_4$ commutes with $-I$, $\mathbf{S}_4$ acts on $\mathrm{Fix}(-I)$. It is a curious fact that the only nontrivial irreducible representation of $\mathbf{S}_4$ acting on the polarity pairs of muscle groups is the standard action of the rotation group of the cube acting on $\mathbf{R}^3$.

The Physiological Role of the Muscle Groups

Finally, for purposes of interpretation, we adopt a caricature of the anatomy of the muscle groups, illustrated in Figure 18. Here we assume that the principal effect of a muscle group being activated is to pull the head in the indicated direction. Six muscle groups effectively form a 'hexagon' and their effect is to tilt the head in various directions. The other two, LA and RA, rotate the head about the vertical axis (as sensed by the horizontal canals lh, rh). There is some redundancy here: the hexagon includes three pairs of muscle groups but the three associated directions are linearly dependent. However, the use of six muscles makes the head position more stable, so there may be physiological reasons for this redundancy. McCollum and Boyle [29] call this hexagon the *central dial*.

It remains to be seen whether the octahedral symmetry that is present in the canal-neck projection of the vestibular system can be used to shed light on some of the functions of that system. However, symmetry is rarely an accident.

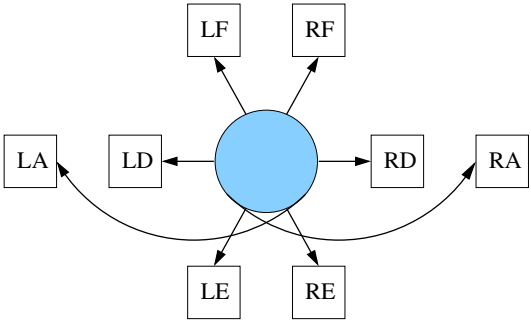


Figure 18: Caricature of effect of activation of muscle groups.

Acknowledgments

There are many people whose ideas concerning the uses of symmetry in neuroscience were used in this report, and I would like to thank each of them: Robert Boyle, Paul Bressloff, Luciano Buono, Jim Collins, Bard Ermentrout, Gin McCollum, LieJune Shiau, Peter Thomas and Matt Wiener, and in particular, Jack Cowan and Ian Stewart. This work was supported in part by NSF Grant DMS-0244529 and in part by the Newton Institute.

References

[1] G.G. Blasdel. Orientation selectivity, preference, and continuity in monkey striate cortex, *J. Neurosci.* **12** (1992) 3139–3161.

[2] J. Blaszczyk and C. Dobrzecka. Alteration in the pattern of locomotion following a partial movement restraint in puppies. *Acta. Neuro. Exp.* **49** (1989) 39–46.

[3] W.H. Bosking, Y. Zhang, B. Schofield, and D. Fitzpatrick. Orientation selectivity and the arrangement of horizontal connections in tree shrew striate cortex, *J. Neurosci.* **17** 6 (1997) 2112–2127.

[4] P.C. Bressloff, J.D. Cowan, M. Golubitsky, and P.J. Thomas. Scalar and pseudoscalar bifurcations motivated by pattern formation on the visual cortex, *Nonlinearity* **14** (2001) 739–775.

[5] P.C. Bressloff, J.D. Cowan, M. Golubitsky, P.J. Thomas, and M.C. Wiener. Geometric visual hallucinations, Euclidean symmetry, and the functional architecture of striate cortex, *Phil. Trans. Royal Soc. London B* **356** (2001) 299–330.

[6] P.C. Bressloff, J.D. Cowan, M. Golubitsky, P.J. Thomas and M.C. Wiener. What geometric visual hallucinations tell us about the visual cortex, *Neural Computation* **14** (2002) 473–491.

[7] P.L. Buono and M. Golubitsky. Models of central pattern generators for quadruped locomotion: I. primary gaits. *J. Math. Biol.* **42** No. 4 (2001) 291–326.

[8] J. Clottes and D. Lewis-Williams. *The Shamans of Prehistory: Trance and Magic in the Painted Caves*, Abrams, New York, 1998.

[9] J.J. Collins and I. Stewart. Hexapodal gaits and coupled nonlinear oscillator models, *Biol. Cybern.* **68** (1993) 287–298.

[10] J.J. Collins and I. Stewart. Coupled nonlinear oscillators and the symmetries of animal gaits, *J. Nonlin. Sci.* **3** (1993) 349–392.

[11] J.D. Cowan. Some remarks on channel bandwidths for visual contrast detection. *Neurosciences Research Program Bull.* **15** (1977) 492–517.

[12] G.B. Ermentrout and J.D. Cowan. A mathematical theory of visual hallucinations, *Biol. Cybern.* **34** (1979) 137–150.

[13] U. Eysel. Turning a corner in vision research, *Nature* **399** (1999) 641–644.

[14] P.P. Gambaryan. *How Mammals Run: Anatomical Adaptations*, Wiley, New York 1974.

[15] C.D. Gilbert. Horizontal integration and cortical dynamics, *Neuron* **9** (1992) 1–13.

[16] M. Golubitsky, L-J. Shiau and A. Torok. Bifurcation on the visual cortex with weakly anisotropic lateral coupling. *SIAM J. Appl. Dynam. Sys.* **2** (2003) 97–143.

[17] M. Golubitsky and I. Stewart. *The Symmetry Perspective: From Equilibrium to Chaos in Phase Space and Physical Space*. Progress in Mathematics **200**, Birkhäuser, Basel, 2002.

[18] M. Golubitsky and I. Stewart. Nonlinear dynamics of networks: the groupoid formalism. *Bull. Amer. Math. Soc.*. To appear.

[19] M. Golubitsky, I. Stewart, P.-L. Buono, and J.J. Collins. A modular network for legged locomotion, *Physica D* **115** (1998) 56–72.

[20] M. Golubitsky, I. Stewart, P.-L. Buono, and J.J. Collins. Symmetry in locomotor central pattern generators and animal gaits, *Nature* **401** (1999) 693–695.

[21] M. Golubitsky, I. Stewart, and D.G. Schaeffer. *Singularities and Groups in Bifurcation Theory II*, Applied Mathematics Sciences, **69**, Springer-Verlag, New York, 1988.

[22] M. Golubitsky, I. Stewart, and L.J. Shiau. Spatio-temporal symmetries in the disynaptic canal-neck projection. In preparation.

[23] M. Golubitsky, I. Stewart, and A. Török. Patterns of synchrony in coupled cell networks with multiple arrows, *SIAM J. Appl. Dynam. Sys.* **4**(1) (2005) 78–100.

[24] S. Grillner, D. Parker, and A.J. El Manira. Vertebrate locomotion—a lamprey perspective, *Ann. New York Acad. Sci.* **860** (1998) 1–18.

[25] S. Grillner and P. Wallén. Central pattern generators for locomotion, with special reference to vertebrates, *Ann. Rev. Neurosci* **8** (1985) 233–261.

[26] D.H. Hubel and T.N. Wiesel. Sequence regularity and geometry of orientation columns in the monkey striate cortex, *J. Comp. Neurol.* **158** No. 3 (1974) 267–294; Uniformity of monkey striate cortex: a parallel relationship between field size, scatter, and magnification factor, *J. Comp. Neurol.* **158** No. 3 (1974) 295–306; Ordered arrangement of orientation columns in monkeys lacking visual experience, *J. Comp. Neurol.* **158** No. 3 (1974) 307–318.

[27] H. Klüver. *Mescal and Mechanisms of Hallucinations*. University of Chicago Press, Chicago, 1966.

[28] N. Kopell and G.B. Ermentrout. Coupled oscillators and the design of central pattern generators, *Math. Biosci.* **90** (1988) 87–109.

[29] G. McCollum and R. Boyle. Rotations in a vertebrate setting: evaluation of the symmetry group of the disynaptic canal-neck projection. *Biol. Cybern.* **90** (2004) 203–217.

[30] A. Patterson. *Rock Art Symbols of the Greater Southwest*, Johnson Books, Boulder CO, 1992, 185.

[31] E. Schwartz. Spatial mapping in the primate sensory projection: analytic structure and relevance to projection, *Biol. Cybernetics* **25** (1977) 181–194.

[32] G. Schöner, W.Y. Jiang, and J.A. Kelso. A synergetic theory of quadrupedal gaits and gait transitions, *J. Theor. Biol.* **142** No 3 (1990) 359–391.

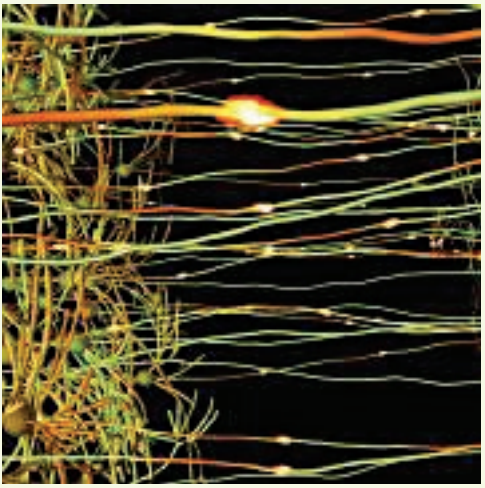
[33] R.K. Siegel and M.E. Jarvik. Drug–induced hallucinations in animals and man. In: *Hallucinations: Behavior, Experience and Theory*, (R.K. Siegel and L.J. West, eds.) Wiley, New York, 1975, 81–161.

[34] L.C. Sincich and G.G. Blasdel. Oriented axon projections in primary visual cortex of the monkey. *J. Neurosci.* **21** (2001) 4416–4426.

[35] I. Stewart, M. Golubitsky, and M. Pivato. Patterns of synchrony in coupled cell networks. *SIAM J. Appl. Dynam. Sys* **2** (2003) 609–646. [DOI: 10.1137/S1111111103419896]

[36] T. Vilis. The physiology of the senses, Lecture 10: Balance, www.med.uwo.ca/physiology/courses/sensesweb

[37] H.R. Wilson and J.D. Cowan. Excitatory and inhibitory interactions in localized populations of model neurons, *Biophys. J.* **12** (1972) 1–24.



2006 BULLETIN CURRENT EVENTS COMMITTEE

David Eisenbud, Chair
Dan Freed
Susan Friedlander
Andrew Granville
John Morgan
Hugo Rossi
Michael Singer
Bryna Kra
László Lovász
Alejandro Adem
Edward Frenkel

Produced by Raquel Storti
American Mathematical Society