

# Dark Matter in the Universe

Martin J. Rees, Cambridge

(Cambridge University Institute of Astronomy)

It is a prime goal of cosmology to understand how the Universe has evolved from an initial dense fireball about  $10^{10}$  years ago to its present state, where galaxies, grouped into clusters, dominate the large-scale cosmic scene.

An impediment to understanding galaxies and their spatial distribution is that the stars and gas we observe may be little more than a tracer for the material that is dynamically dominant. The evidence for 'dark matter' dates back more than fifty years, to the time when Fritz Zwicky realised that the aggregate mass of the observed galaxies in the Coma cluster was inadequate, by a wide margin, to prevent the system from flying apart. The case is now quite compelling (for a recent review <sup>1</sup>), see the proceedings of IAU Symposium 117); what the dark matter consists of is, however, still a mystery.

The masses inferred from *relative motions* of galaxies in apparently-bound groups and clusters exceed by a factor about 10 those inferred from the *internal dynamics* of the luminous parts of galaxies. This apparent discrepancy could be resolved if galaxies were embedded in extensive dark 'halos'. This hypothesis can be checked in some edge-on disc galaxies, where emission from gas can be observed at radii far exceeding the extent of the conspicuous stellar disc. The mass of this gas is itself negligible, but rotation velocities derived from its spectral lines do not fall off as  $r^{-1/2}$ , as would be expected if the gas were orbiting a

mass distribution concentrated at much smaller radii. Instead, the velocity remains almost constant, implying that  $M(< r) \propto r$ , out to  $\approx 80$  kpc in some cases (see Fig. 1).

Estimates of the masses of galaxy clusters come from the virial theorem: the gravitational binding energy must be half the internal kinetic energy if they are to be in equilibrium. This method is now complemented by X-ray studies of thermal emission from hot gas in clusters, which probe the depth of the gravitational potential well. In well-studied clusters which appear to have reached virial equilibrium, the mass-to-light ratio  $M/L$  is typically 200 solar units. The dark matter must mainly be in some unknown form — neither ordinary stars nor the gas that emits X-rays suffices. The data are summarised in Fig. 2.

All things considered, the existence of dark matter is quite unsurprising — there are all too many forms it could take, and the aim of observers and theorists must be to narrow down the options.

The present Hubble timescale  $t_H$  is still uncertain, so in quoting numerical values I shall follow a widespread convention and introduce a useful quantity  $h = (3 \times 10^{17} \text{ s}/t_H)$ . The experts advocate values of  $h$  in the range 0.5–1. I shall also refer to average densities of matter in terms of  $\Omega$ , fractions of the critical density  $\rho_{\text{crit}} (\text{kg}/\text{m}^3) = (8\pi G t_H^2/3)^{-1}$  above which the Universe would just close. This density were it all in baryonic form would correspond to a particle den-

sity of  $11 h^2/\text{m}^3$  so the real average density  $n_b$  becomes  $n_b = 11 \Omega_b h^2/\text{m}^3$ . An important number which is perhaps of fundamental significance is the ratio of the baryon density to the density  $n_\gamma$  of photons in the microwave background, apparently a black body with  $T \approx 2.7$  K. This is then  $3 \times 10^{-8} (\Omega_b h^2)$ ; it is a number that grand unified theories (GUTs) must attempt to explain in terms of an inherent favouritism of matter over antimatter at the ultra-early epoch when symmetry-breaking occurred.

A direct lower limit on  $\Omega_b$  of order  $10^{-2}$  can be set from the observed mean density of baryons in conspicuously 'luminous' form (visible stars and gas in galaxies, and intergalactic gas revealed by its X-ray emission), implying a baryon-photon ratio of not less than  $3 \times 10^{-10} h^2$ .

The inferred dark matter in the halos of individual galaxies and in clusters of galaxies apparently contributes a fraction  $\Omega = 0.1\text{--}0.2$  of the critical cosmological density. Its smoother and less clumped distribution suggests that it underwent less dissipation during the processes of galaxy formation than the luminous stars and gas.

Three strands of evidence could eventually pin down what the dark matter in halos and clusters of galaxies really is.

(i) *Particle physics*. When our theories of high-energy physics become less speculative, and we can calculate how many particles of each species (with known mass) should have survived as relics of the big bang.

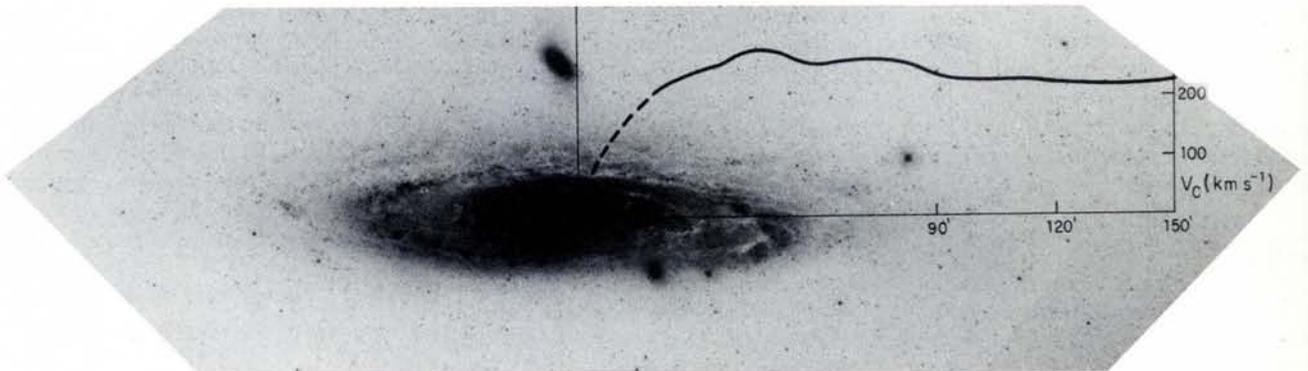


Fig. 1 — The Andromeda Galaxy M31 is shown with, superimposed on it, the rotation velocity of neutral hydrogen, inferred from 21 cm line radio studies. The rotation curve remains 'flat' even beyond the optical outer limits of the galaxy, implying that the outlying gas is 'feeling' the gravitational field of dark matter around the galaxy. (Courtesy of Morton Roberts.)

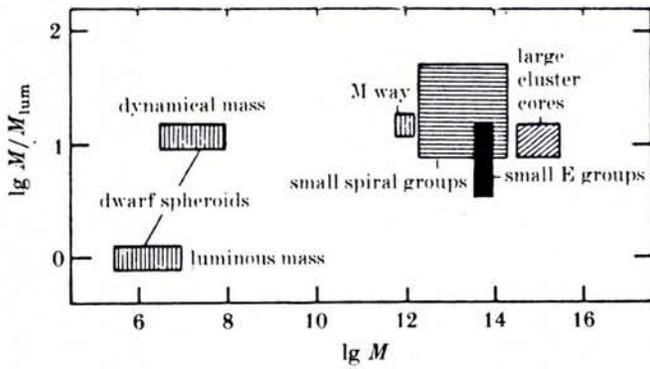


Fig. 2 — The ratio of dark to luminous matter (plotted as a log to base 10) is about 10: 1 on scales larger than individual galaxies. The luminous matter is baryonic, but the dark matter need not be (from Ref. 2).

(ii) *Cosmogonic models.* Evidence on the clustering scales of galaxies, the density profiles in halos, etc. can be compared with the outcome of simulations of gravitational aggregation to see e.g. if the dark matter is 'cold' or not.

(iii) *Direct detection.* The individual entities that our galactic halo is composed of may reveal themselves by astronomical observations, or by direct Earth-based experiments.

### Baryonic Dark Matter

The baryons observed in galaxies and gas, contributing  $\Omega_b \approx 0.01$ , could in principle be dynamically a quite unimportant fraction of all baryons in the Universe. Most of the initial baryons might have condensed into a population of stars that were either pregalactic, or else formed during the initial collapse phase of protogalaxies: could these stars or their remnants, contribute to the unseen mass? Ideally one would like to be able to calculate what happened when the first gravitationally bound clouds condensed from primordial material — to decide whether they form one (or a few) very massive objects (VMOs), or whether fragmentation proceeds efficiently down to low-mass stars. Our poor understanding of what determines the characteristic masses of stars forming now (in, for instance, the Orion nebula) gives us little confidence that we can calculate the nature of stars born in an environment very different from our present-day galaxy.

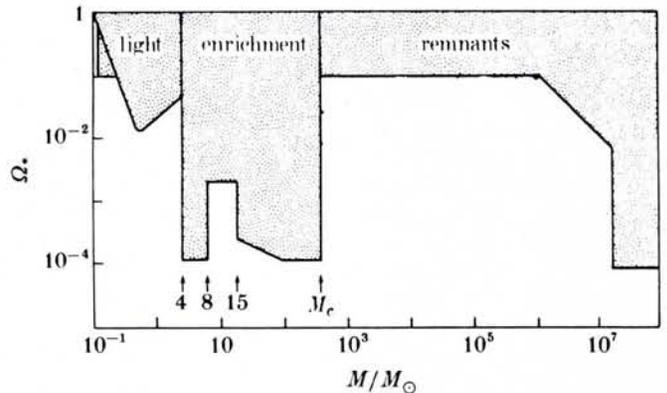
Fig. 3 shows various constraints on the fraction  $\Omega$ , of the critical cosmological density that can be contributed by first-generation stars (or their remnants) in different mass ranges. There are dynamical constraints on the number of  $> 10^6 M_\odot$  black holes in galactic halos, because such massive bodies would have sedimented inwards via dynamical drag on the ordinary stars. The requirements not to overproduce heavy elements constrains the number of remnants of ordinary heavy stars which end their lives by exploding as supernovae. Stars in the mass range  $0.1-1 M_\odot$  would

still be shining after  $\approx 10^{10}$  a, producing too much background light. The only possible baryonic options for dark matter are therefore either 'Jupiters' of below  $0.1 M_\odot$ , or the remnants of very massive objects (VMOs) in the mass range from a few hundred to  $10^6 M_\odot$ , which could have formed at an early time. (These latter objects do not necessarily eject any material processed beyond helium, and leave black hole remnants.)

One way of discriminating between the 'Jupiter' and 'VMO' options is by searching for evidence of gravitational lensing, which occurs whenever a compact mass lies very close to the line of sight to a background source. There is, perhaps surprisingly, much more chance of detecting lensing due to objects in the halos of galaxies (or in clusters) half-way out to the Hubble radius than of seeing a similar effect in our own Galactic Halo. The probability that a compact source (e.g. a distant quasar) is sufficiently closely lined up with an object along the line of sight for lensing to occur is of order  $\Omega_{lens}$ , independent of the individual lens mass involved.

The angular scale,  $\theta_l$ , of the lens image is, however, a diagnostic of the mass and for  $M_l > 10^5 M_\odot$ , very long baseline radio interferometers provide adequate resolution. Results to date suggest that we can probably exclude a major contribution to  $\Omega$  from  $10^6 M_\odot$  objects. For  $M < 0.1 M_\odot$  ('Jupiters')  $\theta_l$  is less than  $(10^{-6})''$ . This cannot be directly resolved by any present day technique,

Fig. 3 — This diagram, adapted and simplified from Ref. 3, shows some constraints on the fraction  $\Omega$  of the cosmological critical density that can be contributed by stars or stellar remnants in various mass ranges.



and will have to await the deployment of adequate optical interferometers in space. There is nevertheless a genuine prospect of detecting lensing of this kind through their transverse motion.

### Constraints on $\Omega_b$ from Primordial Nucleosynthesis

If we accept the concept of light element synthesis in the big bang (occurring when  $kT$  is in the range  $0.1-1$  MeV) then an important constraint on the total baryonic density  $\Omega_b$  comes from deuterium; this is an intermediate product in helium synthesis, the amount that survives being a steeply decreasing function of  $\Omega_b$ . Only if  $\Omega_b h^2 \leq 0.04$  can enough deuterium be produced in a standard hot big bang to provide the observed abundance.

Because the relevant parameter in primordial nucleosynthesis is  $n_b/n_\gamma \sim \Omega_b h^2$ , more precise comparisons of models with observations must await a firmer value of the Hubble constant. If  $h = 1$  (corresponding to a Hubble time of  $10^{10}$  years) then the simplest inference would be that most of the dynamically detected dark matter — both in the halos of individual galaxies and in clusters and groups — was non-baryonic; but if  $h = 0.5$  (corresponding to a Hubble time of  $2 \times 10^{10}$  years), some dark matter — maybe that in halos, if not in intergalactic space — could be baryonic, although only enough to contribute  $\Omega_b \approx 0.1$ . The strength of this limit (which of course precludes  $\Omega_b = 1$  for any reasonable Hubble constant) depends on how compelling the case for 'standard' big bang nucleosynthesis actually is.

### Non-baryonic Dark Matter

If neutrinos have negligible rest mass, the present density that is expected for relic neutrinos from the big bang is  $n_\nu = 1.1 \times 10^8/m^3$  for each two-component species, comparable with the photon density. This conclusion holds for non-zero masses, provided that  $m_\nu c^2$  is far below the thermal energy (5 MeV) at

which neutrinos decoupled from other species and that the neutrinos are stable for the Hubble time. Comparisons with the baryon density shows that neutrinos outnumber baryons by such a big factor that they can be dynamically dominant over baryons even if their masses are only a few electron volts. In fact, a single species of neutrino would yield a contribution to  $\Omega$  of  $\Omega_{\nu} = 0.01 h^{-3} (m_{\nu})_{\text{eV}}$ ; so if  $h = 0.5$ , only 25 eV is sufficient to provide the critical density.

More than 15 years ago, it was conjectured that neutrinos could provide the 'unseen' mass in galactic halos and clusters. At that time the suggestion was not followed up very extensively; but in the 1980s physicists became more open-minded about the possibility of non-zero neutrino masses. A change of theoretical attitude, coupled with experimental claims that  $m_{\nu_e} \approx 36$  eV stimulated astrophysicists to explore scenarios for galaxy formation in which neutrino clustering and diffusion play a key role. More recently, other kinds of non-baryonic matter, such as supersymmetric particles and axions, have also been considered.

Neutrinos have an 'edge' over other non-baryonic candidates in that they are at least known to exist. New evidence on neutrino masses comes from analysis of the burst from the recent supernova 1987A in the Magellanic Clouds. In traversing the distance of 150,000 light years from the supernova, a neutrino of rest mass  $m_{\nu}$  and energy  $E$  would be delayed, in comparison with a zero rest mass particle, by  $0.025 m_{\nu}^2 (10 \text{ MeV}/E)^2$ . The sharpness of the measured burst implies an upper limit of 10–20 eV on the mass of the electron neutrino. Some authors have suggested that the data offer positive evidence for a mass around 4 eV, but such claims are controversial because we do not know the precise time-structure of the catastrophic stellar implosion that triggered the outburst.

#### Experimental Detection of WIMPs

Non-baryonic matter, if dynamically dominant, has an important influence on the cosmogonic process. In so far as we come to understand galaxy formation, we may therefore acquire circumstantial evidence about whether such matter exists and what form it might take. However, it turns out that one may be able to detect such particles (even in the laboratory) if they constitute the mass of our own galactic halo.

Weakly interacting massive ( $> 1$  GeV) particles (WIMPs) would have cross-sections of order  $10^{-36} \text{ cm}^2$  for interactions with nucleons. Such particles, scattering elastically off nucleons in the

Sun, would lose energy *via* the recoil, and could thereby become trapped. Despite being vastly outnumbered by ordinary nuclei, they could then contribute to energy transport in the solar core, because their mean free path is so long; the central temperature would then be slightly lower, with consequent observational consequences such as an alteration in the frequency of some modes of solar oscillation, and a reduction in the  $^8\text{B}$  neutrino flux. The WIMP hypothesis could conceivably, therefore, account simultaneously for two enigmas — the halo dark matter, and the anomalously low solar neutrino output.

If our galactic halo were indeed comprised of WIMPs, then their density near the Earth would be  $\approx 10^5 (m_{\text{GeV}})^{-1} \text{ m}^{-3}$ , and their typical velocities  $\approx 300$  km/s. There is a genuine prospect of detecting this background in the laboratory<sup>4)</sup>. Whenever one of these particles collided with a nucleus, there would be a recoil, with associated energy in the keV range. This energy deposition could be detected if it occurred in a very cold solid with low heat capacity  $\propto T^3$ . The detection could be achieved by maintaining an array of superconducting grains just below the transition temperature: the heat would then raise the temperature of one grain above the critical value, thereby allowing magnetic flux to penetrate. Alternatively, the photons could be detected before thermalisation, or the thermal pulse could be detected directly.

If the dark matter were in so-called 'invisible axions', there would be a chance of detecting these also *via* conversion into photons on interaction with matter or strong magnetic fields. The prospects here are perhaps rather less promising, because the photons would be in a narrow energy band, depending on the poorly known axion mass, so a broad range of photon energies (spanning the microwave, millimetre and infrared bands) would need to be searched<sup>4)</sup>.

Ingenuous schemes for detecting a halo population of exotic particles seem among the most worthwhile and exciting high-risk experiments in physics or astrophysics — potentially as important as those that led to the discovery of the cosmic microwave background in the 1960s. A null result would surprise nobody; on the other hand, such experiments could reveal new supersymmetric particles (or axions, as the case may be), as well as determining what 90 per cent of our Universe consists of. (Because the detection is sensitive to velocity, they would even reveal the halo's velocity dispersion and rotation. The mean velocity of halo particles relative to the

detector would have an annual variation, because of the Earth's motion around the Sun. Such an annual modulation, with an amplitude of a few per cent and a peak in June, would be an unambiguous signature discriminating against spurious background.)

If the baryon-photon ratio could be calculated in terms of a GUT this would determine  $\Omega_b$ . If  $\Omega_b \neq 1$ , then a strictly flat Universe would require some non-baryonic contribution. Of course, one may eventually have theoretical knowledge of the rest masses of all other relevant particles; such information, in conjunction with knowledge of  $n_b/n_\gamma$ , would determine their contribution to  $\Omega$  also. Looked at from this point of view, it perhaps seems coincidental that non-baryonic matter should dominate, but only by an order of magnitude rather than a vastly larger factor.

#### A Flat ( $\Omega = 1$ ) Universe?

In an influential review paper published in 1974 Gott *et al.*<sup>5)</sup> summarized the evidence bearing on  $\Omega$ . They concluded that the dynamical evidence favoured a value 0.1–0.2, and noted that if the matter were all baryonic, the lower end of this range was compatible with the value favoured by standard big bang nucleosynthesis (for a Hubble time  $t_H = 2 \times 10^{10}$  a, a value consistent with the ages of globular clusters, *etc.*). Much new evidence has accumulated during the last decade, especially on cluster dynamics and element abundances; and some relevant theoretical issues have been refined and elaborated. But, if we were to update Gott *et al.*'s discussion, their net conclusion would not change much.

There has, however, been a marked change in theorists' attitudes. This is partly because non-baryonic matter is now taken much more seriously, and seems in some ways almost a natural expectation. But the main element in the discussion is the concept of an inflationary Universe (an early phase of exponential growth): this is so appealing, and resolves some well-known and stubborn cosmological paradoxes in such a natural way, that it instils a strong prejudice in favour of  $\Omega = 1$ .

#### Inferring $\Omega$ from the Cosmic Deceleration

In the simplest cosmological models  $\Omega$  is directly proportional to the deceleration of the cosmological expansion. Programmes attempting to measure this deceleration have been pursued throughout the past 25 years, but there is now not much optimism that they can

succeed until we understand better how galaxies evolve.

There is a spread in the properties of galaxies, just as there is among individual stars. One needs, therefore, to find some class of galaxies which serve as 'standard candles', so that their distance can be inferred from their apparent brightness. Otherwise there will be a big scatter in the magnitude-redshift plot. The brightest galaxies in clusters are the best candidates, but even they display a 30 per cent scatter. Another problem is that galaxies become almost undetectably faint, even when observed with the largest telescopes, when they are far enough away that the effects of the deceleration should really show up.

But the worst problems stem from galactic evolution. The galaxies seen at large distances are systematically younger than nearby ones. Even if a certain class of galaxy provided precise standard candles at the present, one needs to know how each candle changes as it burns. The galaxies that are crucial for cosmological tests are those whose light has been journeying for at least 5 Ga towards us, which are therefore being seen at only about half their present age.

There are two aspects to the evolutionary correction. In a younger elliptical galaxy many stars would be shining which by now have died, and the present stars would be seen at an earlier stage in their evolution. This alters the brightness and the colour of the galaxy, the trend being that a younger galaxy would appear somewhat brighter than its present-day counterpart. But there is a secondary evolutionary correction stemming from the fact that a galaxy is not a self-contained isolated system. We can see many instances where galaxies seem to be colliding and merging with each other; and in rich clusters the large central galaxies may be cannibalising their smaller neighbours. (In a few billion years this may, incidentally, happen in our own local group. The Andromeda galaxy is falling towards our own Milky Way, and there may be a collision between these two large disc galaxies, the likely remnant being a bloated amorphous 'star pile' resembling an elliptical galaxy.) Many big galaxies, particularly those in the centres of clusters, may result from such mergers, traces of disturbance having by now been erased. This process would obviously result in big galaxies, having been, on average, fainter in the past. There are thus two evolutionary effects, both uncertain (and with opposite sign), either of which could be large enough to mask the geo-

metrical difference between a Universe with  $\Omega = 1$  and one with  $\Omega = 0.2$ .

#### 'Biased' Galaxy Formation

The dynamical evidence from clusters and from galactic halos does not offer evidence for any value of  $\Omega$  higher than  $\approx 0.2$ . If  $\Omega$  is indeed unity, then 80 per cent of the mass is unaccounted for even by dynamical considerations: it is not just the light, but evidence for gravitating matter itself, which is 'missing'. This raises the interesting question of whether the dynamical evidence is nonetheless compatible with  $\Omega = 1$ .

If  $\Omega = 1$ , then the dominant mass must not participate fully in the observed clustering: galaxies must be more 'clumped' than mass in general, so that their spatial distribution enhances and exaggerates the inhomogeneity on large scales. This requires some-kind of 'bias' in the formation of galaxies<sup>6</sup>). There are three ways this would come about:

- (i) The entire Universe may be pervaded by a uniform component of 'missing mass' of a different nature from the clustered dark matter, contributing  $\Omega \approx 0.8$ .
- (ii) There may be one important kind of dark matter, but the baryonic component may be segregated from it on scales up to  $\sim 30h^{-1}$  Mpc, so that galaxy formation occurs only in certain regions. This could result from large-scale fluctuations in the initial baryon/photon ratio, from gas dissipation within superclusters of collisionless dark matter (e.g. neutrinos), or from very energetic winds or blast waves pushing the gas over large distances.

(iii) Less extravagant in energy is the possibility that the large-scale baryon distribution does trace dark matter on all scales  $\geq 1 h^{-1}$  Mpc, but the efficiency with which baryons turn into luminous galaxies is modulated by large-scale environmental effects.

#### Conclusions

At least 90 per cent of the mass of the Universe is 'unaccounted for'. It could in principle take many forms — from  $10^{-5}$  eV axions to  $10^6 M_{\odot}$  black holes — but there are fortunately real hopes of narrowing down the range of options. Tackling this task, and the related question of whether the Universe has the 'critical' density indicated by some cosmological theories, will necessitate advances on a broad front — in experimental and theoretical physics, and in elucidating how galaxies form and evolve.

#### REFERENCES

The following references are reviews that can serve as an entry to the extensive literature:

1. *Proceedings of IAU Symposium 117 on Dark Matter*, eds. J. Kormendy and J. Knapp (Reidel, Dordrecht) 1987.
2. Blumenthal G., Faber S.M., Primack J. and Rees M.J., *Nature* **311** (1984) 517.
3. Carr B.J., Bond J.R. and Arnett W.D., *Astrophys. J.* **227** (1984) 445.
4. Articles by P.M. Smith and J.R. Primack in *Cosmology, Astronomy and Fundamental Physics: 2nd ESO/CERN Symposium*, eds. G. Setti and L. Van Hove.
5. Gott J.R., Gunn J.E., Schramm D.N. and Tinsley B.M., *Astrophys. J.* **194** (1974) 343.
6. Dekel A. and Rees M.J., *Nature* **326** (1987) 455.

## Leon Rosenfeld Postdoctoral Fellowship Nordita, Copenhagen

Nordita and the Leon Rosenfeld Foundation invite applications for the Leon Rosenfeld Fellowship for a two-year period beginning in or around September 1988. Nordita conducts active research programmes in astrophysics, condensed matter physics, dynamical systems, elementary particle physics, and nuclear physics. The successful applicant would be expected to work in an area of his or her choice at Nordita, interacting with the rest of the staff and fellows and with the scientific community in the Nordic countries. Extra funds will be reserved to support travel and collaboration by the fellow with workers at other institutions. Applicants, who may come from any country, should have received their Ph.D. or equivalent degree within the last few years or expect to receive it before they begin the fellowship. They should send a curriculum vitae and publication list and arrange for three letters of recommendation to be sent no later than December 1, 1987, to

The Director, Nordita

Blegdamsvej 17, DK - 2100 Copenhagen  $\phi$ , Denmark

from whom information about salary and other conditions may also be obtained. The decision on the choice of the successful applicant will be made early in 1988.