Transient optical emission from the error box of the γ-ray burst of 28 February 1997

J. van Paradijs[1,2], P. J. Groot[1], T. Galama[1],
C. Kouveliotou[3,4], R. G. Strom[5,1], J. Telting[5,6],
R. G. M. Rutten[5,6], G. J. Fishman[4], C. A. Meegan[4],
M. Pettini[7], N. Tanvir[8], J. Bloom[8], H. Pedersen[9],
H. U. Nørdgaard-Nielsen[10], M. Linden-Vørnle[10],
J. Melnick[11], G. van der Steene[11], M. Bremer[12],
R. Naber[13], J. Heise[14], J. in 't Zand[14], E. Costa[15],
M. Feroci[15], L. Piro [15], F. Frontera[16], G. Zavattini[16],
L. Nicastro[17], E. Palazzi[17], K. Bennet[18], L. Hanlon[19]
& A. Parmar[18]

For almost a quarter of a century[1], the origin of γ-ray bursts—brief, energetic bursts of high-energy photons—has remained unknown. The detection of a counterpart at another wavelength has long been thought to be a key to understanding the nature of these bursts (see, for example, ref. 2), but intensive searches have not revealed such a counterpart. The distribution and properties of the bursts[3] are explained naturally if they lie at cosmological distances (a few Gpc)[4], but there is a countervailing view that they are relatively local objects[5], perhaps distributed in a very large halo around our Galaxy. Here we report the detection of a transient and fading optical source in the error box associated with the burst GRB970228, less than 21 hours after the burst[6,7]. The optical transient appears to be associated with a faint galaxy[7,8], suggesting that the burst occurred in that galaxy and thus that γ-ray bursts in general lie at cosmological distance.

GRB970228 was detected[9] with the Gamma-ray Burst Monitor[10] on board the Italian–Dutch BeppoSAX satellite[11] on 1997 February 28, UT 02 h 58 min 01 s. The event lasted ∼80 s and reached peak fluxes of $\sim4 \times 10^{-6}$, $\sim6 \times 10^{-6}$ and $\sim10^{-7}\,erg\,cm^{-2}\,s^{-1}$ in the 40–600 keV, 40–1,000 keV and 1.5–7.8 keV ranges, respectively[9,12] (note that the peak flux of 0.23 Crab quoted in ref. 9 is in error). It occurred in the field of view of one of the BeppoSAX Wide Field Cameras (WFCs)[13]. The spectrum of the event is characteristic of classical γ-ray bursts (GRBs)[12]. Its position (about halfway between α Tauri and γ Orionis) was determined with an accuracy of 3′ (radius)[7] at right ascension (RA) 05 h 01 min 57 s, declination (dec.) +11° 46.4′. Application of the long-baseline timing technique[14] to the GRB data obtained with the Ulysses spacecraft, and with the BeppoSAX and the Wind satellites, respectively, constrained this location to be within each of two parallel annuli, with half-widths[15,16] of 31″ (3σ), and 30″ (3σ), respectively, which intersect the WFC error circle (Fig. 1).

Eight hours after the burst occurred, BeppoSAX was reoriented so that the GRB position could be observed with the LECS and MECS detectors[17,18]. A weak X-ray source was then found[19] at RA

05 h 01 min 44 s, dec. +11° 46.7′ (error radius 50″), near the edge of the WFC error circle[9] (Fig. 1). The 2–10 keV (MECS) flux of this source was $2.4 \times 10^{-12}\,erg\,cm^{-2}\,s^{-1}$. The LECS instrument measured a 0.1–10 keV source flux of $(2.6 \pm 0.6) \times 10^{-12}\,erg\,cm^{-2}\,s^{-1}$. The source spectrum was consistent with a power-law model with photon index 2.7, reduced at low energy by a column density N_H of $5.6 \times 10^{21}\,cm^{-2}$. During an observation with the same instruments on March 3, UT 17 h 37 min this flux had decreased by a factor of 20 (ref. 19). With ASCA the X-ray source was detected[20] on 7 March at a 2–10 keV flux of $(0.8 \pm 0.2) \times 10^{-13}\,erg\,cm^{-2}\,s^{-1}$.

On February 28, UT 23 h 48 min, 20.8 hours after the GRB occurred, before we had any knowledge of the X-ray transient, we obtained a V-band and an I-band image (exposure times 300 s each) of the WFC error box with the Prime Focus Camera of the 4.2-m William Herschel Telescope (WHT) on La Palma[21]. The 1,024 × 1,024 pixel CCD frames (pixel size 24 μm, corresponding to 0.421″) cover a 7.2′ × 7.2′ field, well matched to the size of the GRB error box. The limiting magnitudes of the images are $V = 23.7$, and $I = 21.4$. We obtained a second I-band image on March 8, UT 21 h 12 m with the same instrument on the WHT (exposure time 900 s), and a second V-band image on March 8, UT 20 h 42 m with the Isaac Newton Telescope (INT) on La Palma (exposure 2,500 s). Photometric calibration was obtained from images of standard star number 336 and Landolt[22] field 104. The images were reduced using standard bias subtraction and flatfielding.

A comparison of the two image pairs immediately revealed one object with a large brightness variation[6]: it is clearly detected in both the V- and I-band images taken on 28 February, but not in the second pair of images taken on 8 March (Fig. 2). From a comparison with positions of nearby stars that were obtained using the Digitized Sky Survey we find for its location RA 05 h 01 min 46.66 s, dec. +11° 46′ 53.9″ (equinox J2000); this position has an estimated (internal) accuracy of 0.2″. The object is located in the error box defined by the WFC position, the Ulysses/BeppoSAX/Wind annuli, and the transient X-ray source position (Fig. 1).
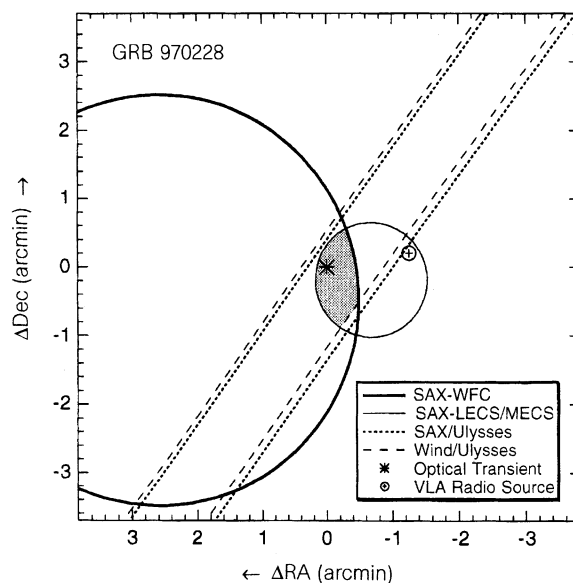
Figure 1 The position of the optical transient, indicated with an asterisk, is shown with respect to the 3′ (radius) WFC location error circle, the 50″ (radius) error circles of the BeppoSAX X-ray transient, and the two annuli obtained from the differences between the times the GRB was detected with Ulysses, and with BeppoSAX and Wind, respectively. The area in common between these error regions is hatched. The coordinates are given in units of arcmin with respect to the position of the optical transient (RA 05 h 01 min 46.665 dec +11° 46′ 53.9″, J2000). The position of an unrelated radio source[49] in the error circle of the X-ray transient is indicated with the square symbol.

Using aperture photometry software we determined the magnitudes of the variable as follows[6]: $V = 21.3 \pm 0.1$, $I = 20.6 \pm 0.1$ on 28 February and $V > 23.6$, $I > 22.2$ on 8 March. The shape of the source in both the 28 February V- and I-band images is consistent with that of the point-spread function, as determined for 15 stars in the same images.

Close to the optical transient is a star, located 2.85″ away at RA 5 h 01 min 46.47 s, dec. $+11° 46' 54.0''$, with $V = 23.1$, $I = 20.5$. A spectrum of this star, taken on March 1, UT 0 h with the ESO 3.6-m telescope using the EFOSC1 spectrograph and the R1000 grating (resolution of 14 Å per pixel), covering the 5,600–11,000 Å region, reveals the presence of TiO bands, which indicate it is an M-type star. With foreground absorption $A_v = 0.4 \pm 0.3$ mag (ref. 23), (substantially smaller than the value inferred from the low-energy cut off in the LECS spectrum), its colour index, $V - I = 2.6$, corresponds to an M2 star[24]. It is most likely to be an M dwarf at a distance of ~3 kpc (an early M-type giant would be located at a distance of ~0.4 Mpc, which we consider much less likely).

Further images were obtained with the Nordic Optical Telescope (NOT, La Palma) on 4 March, with the INT on 9 March, and with the ESO New Technology Telescope (NTT) on 13 March (see Table 1 for a summary). The transient was not detected in these images, which puts a lower limit on its average decay rate (in 4 days) of 0.7 mag per day. The NTT image shows that at the location of the variable object there is an extended object, probably a galaxy[7,8] (Fig. 3); this object is also seen in the INT B- and R-band images. From differential astrometry relative to the nearby M star for both the V- and I-band images, we find that the centres of the optical transient and the galaxy have a relative distance $(0.22 \pm 0.12)''$ (1σ; quadratic addition of the errors in the two relative positions). The relative position of the optical transient did not change by more than 0.2″ between the 28 February V- and I-band images. From the NTT image and the 9 March INT image we measured[7] the galaxy's magnitude to be $R = 23.8 \pm 0.2$ and 24.0 ± 0.2, respectively, consistent with the value $R = 24.0$ reported by Metzger et al.[8], and $B = 25.4 \pm 0.4$ (Table 1).

Known types of optical transient events (novae, supernovae, dwarf novae, flare stars) are unlikely to account for the optical transient for a variety of reasons, such as the amplitude and short timescale of its variability, its colour index, or its inferred distance.

The GRB source is located relatively close to the ecliptic, at latitude $-11°$, and this raises the possibility that the optical transient is an asteroid. However, on 1 March asteroids in the direction of the GRB have proper motions of at least $0.1°$ per day (T. Gehrels, personal communication), which would have led to easily detectable motion ($>2.5''$) during the 600-s total exposure time of our two separate images. On the basis of its proper motion during our two exposures we cannot rule out that the optical transient is a Kuiper belt object. However, its non-detection on other images taken in the week following the GRB, the low surface density (one per several hundred square degrees at 21st magnitude; T. Gehrels, personal communication) of Kuiper belt objects, and their very red colours[25], make such objects a highly unlikely explanation of the optical transient.

The variability of the optical sky above 21st magnitude, on timescales of a few days, has not yet been extensively explored[26,27], and it is therefore not possible to make a firm estimate of the probability that the optical transient we detected is a faint, strongly variable AGN unrelated to the GRB, or has another (unknown) origin unrelated to the GRB. Some information is available from the faint-galaxy monitoring program of Kochanski et al.[28] which covered 2,830 galaxies (down to $B = 24.8$, $R = 23.3$) in a $16' \times 16'$ field during 10 years. They found that near $B = 24$ only ~0.5% of these varied by more than 0.03 mag on a timescale of months to years; none varied by more than a magnitude. Variability of blazars on timescales of minutes to hours does not exceed 0.3 mag; day-to-day variations of a factor of two or more have been observed[50].

Although we cannot firmly exclude that the optical transient in the error box of GRB970228 is caused by some unknown event unrelated to the GRB, the temporal coincidence between the optical and X-ray transients, and their spatial coincidence with the GRB lead us to believe that both the optical transient and the decaying BeppoSAX X-ray source are associated with GRB970228.

Radio observations (at 6 cm wavelength) of the GRB error box[6] made with the Westerbork Radio Synthesis telescope on February 28, UT 23 h 17 min (for 1.2 h; 20.4 h after the GRB, simultaneous with the observations at the WHT), and on March 1, UT 18 h and March 2, UT 18 h (each lasting 12 hours) show that at the position of the optical transient there is then no radio point source with a flux exceeding 1.0, 0.33 and 0.33 mJy, respectively (2σ upper limit).

Some rough spectral information on the optical/X-ray transient can be obtained if we assume that between the two BeppoSAX X-ray observations, made 4 days apart, the X-ray flux of the transient decreased with time since the GRB as a power law. Approximating the spectrum by $F_\nu \propto \nu^\alpha$ we estimate from $F_X \simeq 0.04\,\mu$Jy (interpolated) and $F_V \simeq 10\,\mu$Jy, that $\alpha = -0.7 \pm 0.1$. Extrapolation of this spectrum to the radio region would lead to an expected radio flux density of 10–100 mJy, far exceeding the observed upper limit. This indicates that the X-ray, optical and radio flux densities of the transient cannot be represented by a single power law.

The close positional coincidence between the optical transient and the galaxy suggests that the transient may be located in that galaxy. In an effort to quantify this we adopt a bayesian approach. We consider three disjoint and exhaustive hypotheses for the source
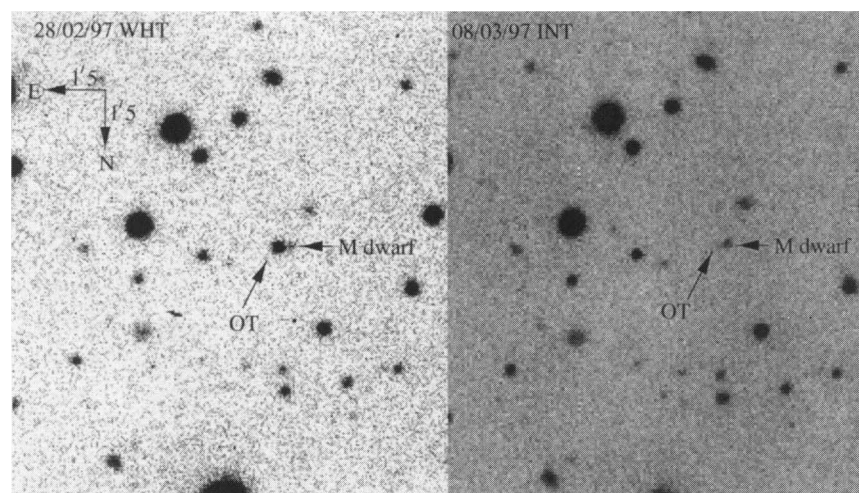
Figure 2 V-band images of a 1.5′ × 1.5′ region of the sky containing the position of the optical transient. The left image was obtained with the WHT on 1997 28 February, UT 23 h 48 min, the right image with the INT on 8 March, UT 20 h 42 min. The optical transient is indicated by 'OT'. The M dwarf, separated from the optical transient by 2.9″, is also indicated.
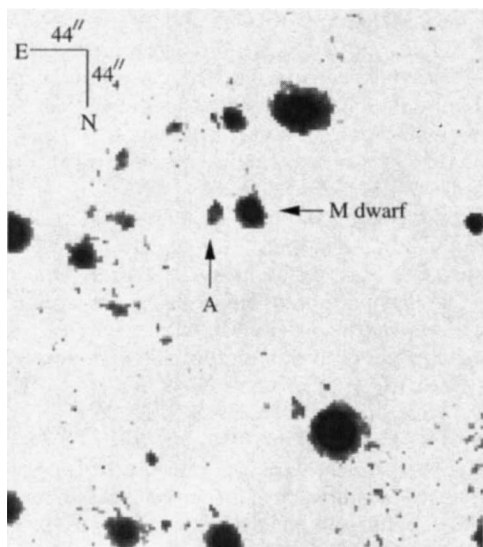
Figure 3 R-band image of a 44″ × 44″ region of the sky containing the position of the optical transient, obtained with the NTT on 1997 13 March UT 0 h. The faint galaxy coincident with the optical transient (A) and the M-dwarf are indicated.

Table 1 Summary of optical observations

Date (1977)	Time (UT)	Telescope	Band	Integration time (s)	Magnitude	Remarks
28 Feb.	23 h 48 min	WHT	V	300	$V = 21.3$	Transient
28 Feb.	23 h 53 min	WHT	I	300	$I = 20.6$	Transient
04 Mar.	20 h 42 min	NOT	V	900	$V > 24.2$	–
08 Mar.	20 h 42 min	INT	V	2,500	$V > 23.6$	–
08 Mar.	21 h 12 min	WHT	I	300	$V > 22.2$	–
09 Mar.	21 h 30 min	INT	R	1,200	$R = 24.0$	Extended
09 Mar.	20 h 30 min	INT	B	2,500	$B = 25.4$	Extended
13 Mar.	0 h 0 min	NTT	R	3,600	$R = 23.8$	Extended

of the optical transient. The hypotheses are: H_c, the optical transient is in the centre of a galaxy; H_g, the optical transient is in a galaxy but not at its centre; and H_n, the optical transient is not in a galaxy.

We assume that there is a single optical transient detected in the field of view of angular area A and that n non-overlapping galaxies are detected in the field. The transient is at distance $r \pm \sigma_r$ from the nearest galaxy, where the error includes the uncertainty in the positions of the centroids of both the galaxy and the transient. The probability density at r under H_c is $P(r \mid H_c) = (2\pi\sigma_r^2 n)^{-1}\exp[-(r^2/2\sigma_r^2)]$. The probability density under H_g depends on the size, shape and inclination of the galaxy and the specifics of the model for the distribution of sources in the galaxy. For simplicity, we assume the probability density to be gaussian with width σ_g. Then $P(r \mid H_g) = (2\pi\sigma_g^2 n)^{-1}\exp[-(r^2/2\sigma_g^2)]$. The probability density under H_n is uniform over the field of view, so $P(r \mid H_n) = A^{-1}$. The posterior probability of each hypothesis H is $P(H \mid r) = kP(H)P(r \mid H)$, where $P(H)$ is the prior probability and the normalization constant k is obtained by the requirement that $P(H_c \mid r) + P(H_g \mid r) + P(H_n \mid r) = 1$.

For the NTT observation we find seven galaxies in $A = (44'')^2$ field, that is, $n = 13$ per arcmin^2, $r = 0.22''$, $\sigma_r = 0.12''$. A reasonable estimate for the galaxy width is $\sigma_g = 1''$. With these values the probability densities at r are $P(r \mid H_c) = 0.294$, $P(r \mid H_g) = 0.022$ and $P(r \mid H_n) = 0.0005$, all in units of arcsec^{-2}. Assuming equal priors $P(H_c) = P(H_g) = P(H_n) = 1/3$, the posterior probabilities are $P(H_c \mid r) = 0.928$, $P(H_g \mid r) = 0.070$ and $P(H_n \mid r) = 0.0016$. The posterior probability for H_g depends sensitively on the assumed σ_g. However, the posterior probability, $P(H_n \mid r)$, that the transient is not associated with a galaxy is in the range 0.09–0.18%, for any assumptions about the size of faint galaxies. Within the range of assumed values for σ_r between 0.08″ and 0.2″ the values of $P(H_n \mid r)$ increase by less than a factor of 3.

The above analysis suggests that the optical transient is related to the faint galaxy, which provides support for the cosmological distance scale for GRBs.

A rough estimate of the expected redshift, z, of the galaxy may be made by assuming that its absolute magnitude is in the range -21 to -16, which covers the bulk of normal galaxies[29]. For an assumed Hubble constant of 60 km s^{-1} Mpc^{-1}, this corresponds to z in the range 0.2–2.

The close proximity of the optical transient to the centre of the faint galaxy, and the presence of relatively bright quasars in the

8 arcmin^2 error box of GRB781119 ($V = 20$)[30–32], and in the 3′ (radius) error box[33] of GRB960720 ($R = 18.8$)[34–36] raise the possibility that GRBs occur preferentially, or exclusively, in or near galactic nuclei.

Searches for an optical counterpart to a GRB have been continually attempted for the past 20 years. Recent reviews and descriptions of serendipitous, rapid follow-up, and delayed searches for optical counterparts of GRBs[36–44] show that these previous searches were generally made a week or longer after the GRB, or they were not as deep ($V < 20$) as the images presented here. The most sensitive rapid follow-up observations so far had delay times (δt) and limiting magnitudes (m) as follows: $\delta t = 1.85$ d, $m < 23$ (ref. 45), and $\delta t = 4.0$ d, $m_B < 22$ (ref. 38).

It was not until the launch of BeppoSAX in 1996 that accurate (several arcmin) locations for GRBs became available within hours of detection, hence facilitating rapid follow-up observations at large ground-based optical telescopes for those bursts which happened to be in the field of the WFC. The continued operation of BeppoSAX and the approval of the High Energy Transient Explorer-2 (HETE-2) mission bode well for great progress in the rapid follow-up observations of GRBs. Also, near-real-time, fully automated optical systems linked to the BATSE-BACODINE[37] system are becoming operational, and their sensitivity is continually improving[46,47].

We expect that X-ray and optical transients associated with GRBs will again be seen (though perhaps not in all cases[48]) in the near future. This could be a turning point in GRB astronomy. Detailed studies (light curves and spectra) of such transients can be expected within a year, and we are optimistic that the distance scale as well as the mechanism behind the enigmatic GRBs are now within reach. *Note added in proof*: After this paper was submitted, an HST observation was made of the optical transient (K. Sahu *et al.*, *IAU Circ. No.* 6606). This observation confirms that the transient is associated with an extended emission region, but seems to exclude that the transient is located at the centre of that region. □

1. Klebesadel, R. W., Strong, I. B. & Olson, R. A. Observations of gamma-ray bursts of cosmic origin. *Astrophys. J.* **182**, L85–L88 (1973).
2. Fishman, G. J. & Meegan, C. A. Gamma-ray bursts. *Annu. Rev. Astron. Astrophys.* **33**, 415–458 (1995).
3. Meegan, C. A. *et al.* Spatial distribution of gamma-ray bursts observed by BATSE. *Nature* **355**, 143–145 (1992).
4. Paczyński, B. Gamma-ray bursters at cosmological distances. *Astrophys. J.* **308**, L43–L46 (1986).
5. Podsiadlowski, Ph., Rees, M. J. & Ruderman, M. Gamma-ray bursts and the structure of the Galactic halo. *Mon. Not. R. Astron. Soc.* **273**, 755–771 (1995).
6. Groot, P. J. *et al. IAU Circ. No.* 6584 (1997).
7. Groot, P. J. *et al. IAU Circ. No.* 6588 (1997).
8. Metzger, M. R. *et al. IAU Circ. No.* 6588 (1997).
9. Costa, E. *et al. IAU Circ. No.* 6572 (1997).
10. Frontera, F. *et al.* The high-energy X-ray experiment PDS on board the SAX satellite. *Adv. Space Res.* **11**, 281–286 (1991).
11. Piro, L., Scarsi, L. & Butler, R. C. SAX: the wideband mission for X-ray Astronomy, *Proc. SPIE* **2517**, 169–181 (1995).
12. Palmer, D. *et al. IAU Circ. No.* 6577 (1997).
13. Jager, R., Heise, J., In't Zand, J. & Brinkman, A. C. Wide field cameras for SAX. *Adv. Space Res.* **13**, 315–318 (1995).
14. Hurley, K. *et al.* The Ulysses supplement to the BATSE 3B Catalog. *AIP Proc.* **384**, 422–426 (1996).
15. Cline, T. L. *et al. IAU Circ. No.* 6593 (1997).
16. Hurley, K. *et al. IAU Circ. No.* 6594 (1997).
17. Parmar, A. N. *et al.* The low-energy concentrator spectrometer on-board the SAX X-Ray astronomy satellite. *Astron. Astrophys.* (in the press).
18. Bonura, A. *et al.* Performance characteristics of the scientific model of the medium energy concentrator spectrometer on board the X-ray astronomy satellite SAX. *Proc. SPIE* **1743**, 510–522 (1992).

19. Costa, E. *et al. IAU Circ. No.* 6576 (1997).
20. Yoshida, A. *et al. IAU Circ. No.* 6593 (1997).
21. Groot, P. *et al. IAU Circ. No.* 6574 (1997).
22. Landolt, A. UBVRI photometric standard stars in the magnitude range 11.5–16.0 around the celestial equator. *Astron. J.* **104**, 340–376 (1992).
23. Hakkila, J., Myers, J. M., Stidham, B. J. & Hartmann, D. H. A computerized model of large-scale visual interstellar extinction. *Astron. J.* (submitted).
24. Johnson, H. L. Astronomical measurements in the infrared. *Annu. Rev. Astron. Astrophys.* **4**, 191–206 (1966).
25. Luu, J. & Jewitt, D. Color diversity among the Centaurs and Kuiper Belt Objects. *Astron. J.* **112**, 2310–2318 (1996).
26. Paczynski, B. *Variable Stars and the Astrophysical Returns of Microlensing Surveys* (ed. Ferlet, R. and Maillard, J.-P.) (Proc. 12th AIP Colloq., in the press).
27. Trevese, D., Pittella, G., Kron, R. G., Koo, D. C. & Bershady, M. A survey for faint variable objects in SA 57. *Astron. J.* **98**, 108–116 (1989).
28. Kochanski, G. P., Tyson, J. A. & Fischer, P. Flickering faint galaxies: few and far between. *Astron J.* (in the press).
29. Schechter, P. An analytic expression for the luminosity function for galaxies. *Astrophys. J.* **203**, 297–306 (1976).
30. Pedersen, H. *et al.* Optical candidates for the 1978 November 19 gamma-ray burst source. *Astrophys. J.* **270**, L43–L47 (1983).
31. Pedersen, H. & Hansen, J. A quasar in the 1978 November 19 gamma-ray burst error box. *Astrophys. J.* (submitted).
32. Boer, M. *et al.* ROSAT detection and high precision localization of X-ray sources in the November 19, 1978 gamma-ray burst error box. *Astrophys. J.* (in the press).
33. In't Zand, J. *et al. IAU Circ. No.* 6969 (1997).
34. Greiner, J. & Heise, J. *IAU Circ. No.* 6570 (1997).
35. Piro, L. *et al. IAU Circ. No.* 6570 (1997).
36. Walsh, D. *et al.* Spectroscopy of 26 QSO candidates from the Jodrell Bank 966-MHz survey. *Mon. Not. R. Astron. Soc.* **211**, 105–109 (1984).
37. Barthelmy, S. *et al.* Progress with the real-time GRB coordinates distribution Network (BACODINE). *AIP Proc.* **384**, 580–584 (1996).
38. McNamara, B. *et al.* Ground-based gamma-ray burst follow-up efforts: results from the first two years of the BATSE/COMPTEL/NMSU Rapid Response Network. *Astrophys. J. Suppl. Ser.* **103**, 173–181 (1996).
39. Vrba, F., Hartmann, D. & Jennings, M. Deep optical counterpart searches of gamma-ray burst localizations. *Astrophys. J.* **446**, 115–149 (1995).
40. Vrba, F. Searches for gamma-ray burst counterparts: current status and future prospects. *AIP Proc.* **384**, 565–574 (1996).
41. Luginbuhl, C., Vrba, F., Hudec, R., Hartmann, D. & Hurley, K. Results from the USNO quiescent optical counterpart search of IPN[3] GRB and optical transient localizations. *AIP Proc.* **384**, 676–679 (1996).
42. McNamara, B. *et al.* Ground-based γ-ray burst follow-up efforts: the first three years of the BATSE/COMPTEL/NMSU γ-Ray Burst Rapid Response Network. *AIP Proc.* **384**, 680–684 (1996).
43. Castro-Tirado, A., Brandt, S., Lund, N. & Guziy, A. Optical follow-up of gamma-ray bursts observed by WATCH. *AIP Proc.* **307**, 404–407 (1994).
44. Klose, S. Search for an optical counterpart of the source of GRB911001. *Astrophys. J.* **446**, 357–360 (1995).
45. Schaefer, B. *et al.* Rapid searches for counterparts of GRB 930131. *Astrophys. J.* **422**, L71–L74 (1994).
46. Park, H. S. *et al.* Limits on Real-time Optical Emission from Gamma-Ray Bursts Measured by the GROCSE Experiment. *Astrophys. J.* (submitted).
47. Lee, B. *et al.* Results from GROCSE: a real-time search for gamma-ray burst optical counterparts. *Astrophys. J.* (submitted).
48. Castro-Tirado, A. *et al. IAU Circ. No.* 6598 (1997).
49. Frail, D. *et al. IAU Circ. No.* 6576 (1997).
50. Miller, H. R. & Noble, J. C. The microvariability of blazars and related AGN. *ASP Conf. Ser.* **110**, 17–29 (1996).

Momentum creation by vortices in superfluid 3He as a model of primordial baryogenesis

T. D. C. Bevan*, A. J. Manninen*†, J. B. Cook*, J. R. Hook*, H. E. Hall*, T. Vachaspati‡ & G. E. Volovik§‖

* *Schuster Laboratory, University of Manchester, Manchester M13 9PL, UK*
‡ *Physics Department, Case Western Reserve University, Cleveland, Ohio 44106, USA*
§ *Low Temperature Laboratory, Helsinki University of Technology, 02150 Espoo, Finland*
‖ *Landau Institute for Theoretical Physics, 117334 Moscow, Russia*

The Universe contains much more matter than antimatter, which is probably the result of processes in the early Universe in which baryon number was not conserved. These processes may have occurred during the electroweak phase transition, when elementary particles first acquired mass[1–4]. It is impossible to study

directly processes relevant to the early Universe, because of the extreme energies involved. One is therefore forced to investigate laboratory systems with analogous phase transitions. Much of the behaviour of superfluid 3He is analogous to that predicted within the standard model of the electroweak interaction[5]. Superfluids and liquid crystals have already been used to investigate cosmic-string production[6–11]; here we describe experiments on 3He that demonstrate the creation of excitation momentum (which we call momentogenesis) by quantized vortices in the superfluid. The underlying physics of this process is similar to that associated with the creation of baryons within cosmic strings, and our results provide quantitative support for this type of baryogenesis.

To explain the creation of matter in the early Universe we begin by recalling the relationship $E^2 = m_0^2 c^4 + p^2 c^2$ between energy E and momentum p for relativistic particles with rest mass m_0 (c is the velocity of light). Dirac realized that the square root of this equation, $E = \pm (m_0^2 c^4 + p^2 c^2)^{\frac{1}{2}}$ produced particles with both positive and negative energy. This led to his famous picture of the vacuum state (the state with no real particles) as one in which all the negative energy states are full and all the positive energy states are empty. A real particle is then created as shown in Fig. 1a by excitation of a particle from the 'Dirac sea' of negative-energy states to a positive-energy state. But such processes create matter and antimatter in equal amounts, as the appearance of a hole in the negative energy states is interpreted as the simultaneous creation of an antiparticle.

The net creation of matter in the form of baryons such as protons and neutrons requires processes in which antiparticles are not simultaneously created and thus baryon number is not conserved. In the 'standard model' the baryon number is classically conserved but can be violated by quantum-mechanical effects known generically as 'chiral anomalies'. The process leading to matter creation is called 'spectral flow' and can be pictured as a process in which particles flow from negative energies to positive energies under the influence of an external force. In this way real observable positive-energy particles can appear without simultaneous creation of antiparticles. Figure 1b illustrates a simple example of spectral flow occurring for massless particles with electric charge q (in units of the charge of the electron, e) moving in a magnetic field. The application of an electric field $\mathbf{E}$ leads to the production of particles from the vacuum at the rate

$$\dot{n} = \partial_\mu j^\mu = \frac{q^2}{4\pi^2} \mathbf{E} \cdot \mathbf{B} \qquad (1)$$

per unit volume, where j is the particle current fourvector; factors of $(e/\hbar)$ have been absorbed in the definition of the electric and magnetic fields. This is an anomaly equation for the production of particles from the vacuum of the type found by Adler[12] and by Bell and Jackiw[13] in the context of neutral pion decay. We see that for particle creation it is necessary to have an asymmetric branch of the dispersion relation $E(p)$ which crosses the axis from negative to positive energy. We call such a branch a zero mode branch; a spectrum of this type was first found for vortex core excitations in a superconductor[14].

Similar zero mode branches exist on a cosmic electroweak string (also known as a Z-string), which is a structure of the Higgs field that may have been produced during the electroweak phase transition. The Higgs field gives the particles mass outside the string core. This field vanishes on the string axis and the fermions (quarks and leptons) occurring in the core of the string behave like massless one-dimensional particles. Spectral flow on a Z-string leading to production of baryons from the vacuum with conservation of electric charge is illustrated in Fig. 1c. Motion of the string across a background electromagnetic field[15] or the de-linking of two linked loops[16,17] provides a mechanism for cosmological baryogenesis[18] and could lead to the presence of antimatter in cosmic rays[19].

We should point out that baryon-number violation is only one

† Present address: Physics Department, University of Jyväskylä, PO Box 35, 40351 Jyväskylä, Finland.