ELSEVIER

An infrared background-TeV gamma-ray crisis?

R.J. Protheroe [a,*], H. Meyer [a,b]

[a] *Department of Physics and Mathematical Physics, The University of Adelaide, Adelaide, 5005 Australia*
[b] *Universität Wuppertal, Fachbereich Physik, GaussStr 20, 42097 Wuppertal, Germany*

Abstract

We consider the implications of the recent determination of the universal infrared background for the propagation of photons up to 20 TeV from the active galaxy Markarian 501 as observed by HEGRA. At 20 TeV the mean free path for photon–photon collisions on the infrared background would be much shorter than the distance to Markarian 501, implying absorption factors of the order of $\exp(-10)$, or greater, and consequently an excessive power output for this active galaxy. Possible solutions of this problem are discussed. © 2000 Published by Elsevier Science B.V.

PACS: 95.30.Cq; 98.70.Rz; 98.70.Vc; 98.80.Cq

1. Introduction

The general form of the spectrum of the universal infrared background radiation was predicted by Primack et al. [1], Malkan and Stecker [2], Dwek et al. [3] and Fall et al. [4]. The origin is currently thought to be due to star formation in the early universe producing starlight predominantly in the 1 micron range which was then re-processed into the 100 micron range due to heating of surrounding dust. Early data from the DIRBE and FIRAS instruments on COBE (see Ref. [3] for a review) and from ISO [5] were mainly upper or lower limits which did not severely constrain the models. New results from these experiments are becoming available due to a more sophisticated treatment of the galactic and solar-system radiation foreground subtraction. In particular, new determinations of the infrared background intensity have very recently become available at 1.25, 2.2 and 3.5 microns [6], 15.0 microns [7], 60 and 100 microns [8], and 140 and 240 microns [9]. In Fig. 1 we show the predictions of Malkan and Stecker [2] together with the new observations and recent lower limits [10] based on infrared galaxy source counts and a model including the effects of an infrared burst phase of ultraluminous infrared galaxies. The observations are in excellent agreement with prediction inasmuch as the expected trend (two peaks and valley) is found, the agreement being particularly good between 1.25 and 15 microns. However, the level of the observed flux in the 60–240 micron range is higher, and in the case that this is a truly universal background it has important consequences both for star formation in the early universe, and for intergalactic absorption of TeV gamma-rays which leads to a crisis in our understanding of high energy phenomena.

Immediately following the discovery of Markarian 421 as a TeV source [11] it became clear that interac-

---

* Corresponding author.

E-mail address: rprother@physics.adelaide.edu.au
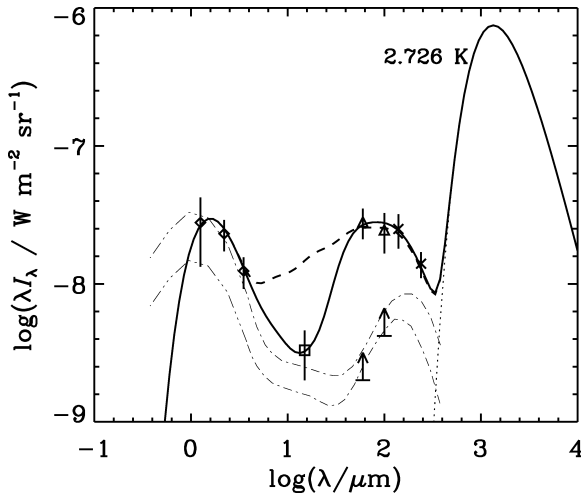(R.J. Protheroe).

Fig. 1. Infrared background radiation field. Recent determinations from DIRBE [6] (diamonds), ISOCAM [7] (square), DIRBE [8] (triangles), and FIRAS [9] (crosses) are compared with the models of Malkan and Stecker [2] (dot-dash curves; the dot-dot-dot dash curves are an extrapolation of the Malkan and Stecker models used in Ref. [20]). The lower limits are from Ref. [10]. In this Letter we model the IR background by the solid curve; thick dashed curve is allowed if the ISOCAM point is considered a lower limit.

tions with a universal infrared background may have a strong influence on the propagation of TeV gamma-rays through intergalactic space [12,13], in the same way that the microwave background affects the propagation of 1000 TeV gamma-rays [14,15]. The observation by several telescopes of a very high level of emission by Markarian 501 during 1997 (see Ref. [16] for a review) showed that the spectrum extended to 10 TeV and beyond. At that time, it was pointed out by Stanev and Franceschini [17] that if the spectrum continued much beyond 10 TeV that we would "have to revise our concepts about the propagation of TeV gamma-rays in the intergalactic space, and that something complicates the process". This has now come about as the analysis of HEGRA data for the whole of 1997 shows that the spectrum of Markarian 501 extends well beyond 10 TeV [18].

2. Implications of the new IR observations for the propagation of Gamma-rays through the universe

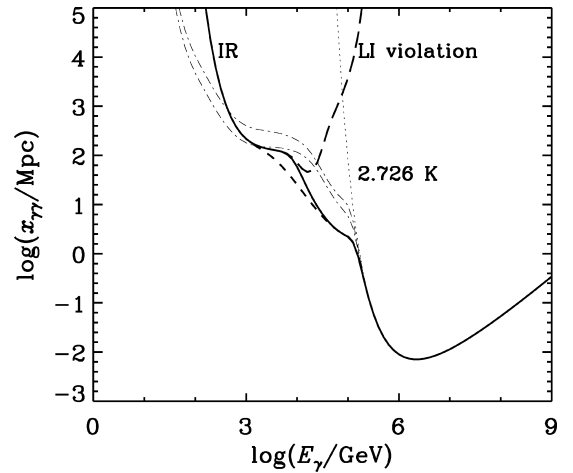Photon–photon pair production is a very elementary process in QED, and can safely be calculated at en-

Fig. 2. Mean free path for photon–photon pair production in the infrared–microwave background radiation. The curves correspond to those in Fig. 1 except that the effect of Lorentz invariance violation discussed in Section 4 is shown by the long dashed curve.

ergies relevant to TeV astronomy. For propagation of energetic particles (gamma-rays) of energy E_γ, mass $m_\gamma(=0)$ and velocity $\beta_\gamma c(=c)$, through isotropic radiation the reciprocal of the mean free path for collisions with photons is given by

$$x_{\gamma\gamma}(E_\gamma)^{-1} = \frac{1}{8E_\gamma{}^2\beta_\gamma} \int\limits_{\varepsilon_{min}}^{\infty} d\varepsilon \frac{n(\varepsilon)}{\varepsilon^2}$$

$$\times \int\limits_{s_{min}}^{s_{max}(\varepsilon,E_\gamma)} ds\, (s - m_\gamma^2 c^4)\sigma(s), \qquad (1)$$

where $n(\varepsilon)$ is the differential photon number density, and $\sigma(s)$ is the total cross section for a centre of momentum frame energy squared given by $s = m_\gamma^2 c^4 + 2\varepsilon E_\gamma(1 - \beta_\gamma\cos\theta)$ where θ is the angle between the directions of the energetic particle (gamma-ray) and soft photon, $s_{min} = (2m_e c^2)^2$, $\varepsilon_{min} = (s_{min} - m_\gamma^2 c^4)/[2E_\gamma(1 + \beta_\gamma)]$, and $s_{max}(\varepsilon, E_\gamma) = m_\gamma^2 c^4 + 2\varepsilon E_\gamma(1 + \beta_\gamma)$. For photon–photon pair production by gamma-rays, we take $m_\gamma = 0$ and $\beta_\gamma = 1$, and the photon–photon pair production cross section from [19].

We have calculated the mean free path for this process for interactions of gamma-rays with the infrared and cosmic microwave backgrounds, and this is shown in Fig. 2. Also shown are results [20] based on
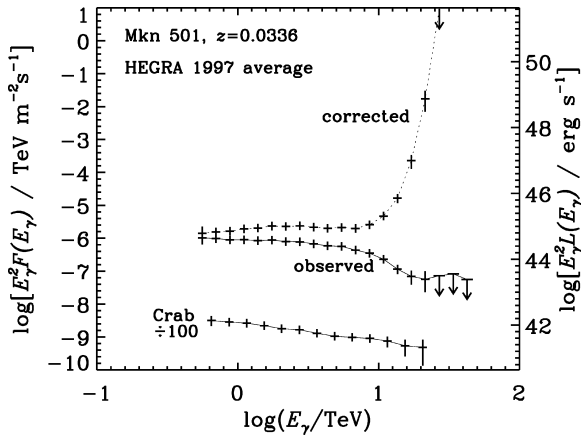
Fig. 3. The time-averaged spectrum of gamma-rays from Markarian 501 observed in 1997 [18] is compared with the spectrum of the Crab Nebula observed in 1997–1998 [21,22]. The spectrum of Markarian 501 after correction for absorption in the infrared background is also shown assuming $H_0 = 65 \text{ km s}^{-1} \text{ Mpc}^{-1}$. The right hand scale shows the luminosity for Markarian 501.

the upper and lower models for the IR background due to Stecker and Malkan [2], extended to shorter wavelengths. We note that above 10 TeV the mean free path using the new IR intensity (thick solid curve) is almost an order of magnitude lower, and this has very important consequences for the HEGRA observations of Markarian 501.

We show in Fig. 3 the time-averaged flux from Markarian 501 observed by HEGRA using the stereoscopic system of telescopes [18]. We note that the spectrum smoothly extends to at least 20 TeV. This observation was confirmed by using the stand-alone telescopes of HEGRA, although with less statistical significance and with rather different systematics [21]. It is important to note that the Crab Nebula, as a steady galactic source, has approximately the same emission level as Markarian 501 in the high state, and has been continuously monitored by the HEGRA telescopes [21,22]. The Crab spectrum is also shown in Fig. 3 for comparison (shifted down by a factor of 100 for clarity). The Crab Nebula spectrum determined extends again up to 20 TeV as a straight power-law with a very similar spectral index, and this suggests reliable detection of photons up to 20 TeV. For the following discussion, an understanding of the absolute energy calibration is of considerable importance. This was investigated in great detail by Monte Carlo methods, and

in addition, checked by absolute measurement of the cosmic ray proton flux using the HEGRA telescope system in about the same energy range as in the Crab and Markarian 501 observations. In this energy range the cosmic ray proton flux is rather well measured by direct methods, and a compilation of data can be found in Ref. [23]. The absolute energy scale of HEGRA was then estimated to have an uncertainty of 15% [24]. In addition, there is data available on Markarian 501 from the Whipple [25] and CAT [26] telescopes, and this is in very good agreement with the HEGRA data but does not reach as high in energy, mainly because of their lower exposure.

We have applied the mean free path in the infrared background radiation plotted in Fig. 2 to "correct" the observed flux from Markarian 501 assuming a Hubble constant of $H_0 = 65 \text{ km s}^{-1} \text{ Mpc}^{-1}$, and this is also shown in Fig. 3. We note that this implies a dramatic (by several orders of magnitude) increase in luminosity at the source (see right hand scale). Even taking into account relativistic beaming, with a typical Doppler factor of 20, this would imply an extraordinarily high luminosity for an active galaxy of the BL Lac class of which Markarian 501 is a member. The large correction factor is a direct consequence of the new IR data at 60 and 100 microns and the implied shift of the "dust peak" to somewhat shorter wavelengths. It is worthwhile pointing out that the correction factor depends on the value of the Hubble constant. If it were larger than that we have assumed, then the distance to the source would be smaller, and the absorption less, but unacceptably high values of H_0 would be required to give a reasonable source luminosity. In fact, the problem may be even more severe if future infrared measurements would fill in the "valley" between 5 and 40 microns giving an IR background such as shown by the short dashed lines in Fig. 1. In this case, the mean free path would be as shown by the short dashed curves in Fig. 2, with the result that correction for absorption would cause an even more dramatic rise of the source spectrum above a few TeV.

3. Immediate consequences

The source spectrum of Markarian 501, i.e., the observed spectrum corrected for propagation from the

source, is in gross disagreement with expectations of the Synchrotron self-Compton (SSC) model routinely used to interpret the observed spectra of active galaxies. In this model, accelerated electrons produce most of the observed radio through X-ray emission as synchrotron radiation, and the same population Compton scatters synchrotron photons to gamma-ray energies. The apparent success of the SSC model depends critically on which intergalactic absorption is used, as the observed sources are at the gamma-ray horizon [27]. Now that the most recent determinations of the diffuse IR background are high, this makes it impossible to fit the gamma-ray observations using the SSC model. For example, Guy et al. [28] took a relatively high level for the IR background, but still a factor of 2 short of recent determination. Petry et al. [29] and Bednarek and Protheroe [20] used the IR background at a level of that predicted by Malkan and Stecker [2], while Sambruna et al. [30] assume a specific shape of the IR background in order to keep the TeV data compatible with the SSC model.

The corrected TeV data (Fig. 3) have a very prominent upturn above 10 TeV which cannot be explained by SSC models. Such an upturn may occur naturally in some proton blazar models [31], particularly those which predict a very flat or nearly monoenergetic spectrum of synchrotron radiating protons [32,33]. However, even with these models it appears impossible to obtain the extreme turn-up in the TeV source spectrum inferred if the new IR data are used to correct for propagation.

In principle the turn up could be a pile-up caused in a pair–Compton cascade during propagation through the IR background [13]. However, for this to be the cause the intergalactic magnetic field would have to be extremely small otherwise the pile-up would be distributed as a halo [34], and any time-variability, such as the observed flaring on time scales of hours [21] would be washed out.

4. Discussion: what is the solution?

Many possible solutions have been mentioned by Finkbeiner et al. [8], and here we provide a quantitative discussion of some of them. The simplest solution to the problem would be if the recent determinations of the universal infrared background were an overes-

timate. This is not completely ruled out, and they are not incompatible with star formation and dust-heating in the early Universe. In a similar manner, a shift in the energy scale of the HEGRA data downwards by a factor of two would have a similar effect. This appears rather extreme, and not very likely. If both the infrared data and the HEGRA TeV data are confirmed, then we may be forced to consider more radical possibilities.

An interesting idea based on known physics was proposed by Harwit, Protheroe and Biermann [35]. It considers a TeV Cherenkov event as being due to a coherent superposition of a number of lower energy photons simultaneously arriving at the top of the atmosphere, and masquerading as a single TeV photon with the sum of the energies. Such Bose–Einstein condensates would suffer losses on a distance scale comparable to that of the mean free path in the infrared radiation of the individual photons of the condensate. This idea can be tested on the basis of extensive simulations, and by detailed comparison of Cherenkov images of gamma ray showers from the Crab nebula and Markarian 501. Such analysis is underway. [1] Even for a Bose–Einstein condensate with, on average, as few as two arriving photons, the effect on the intergalactic absorption of 20 TeV photons from Markarian 501 would be dramatic. We show in Fig. 4 the result of correcting for intergalactic absorption for this case in an approximate way (plotted as asterisks). In calculating this, we assumed for simplicity that each event recorded by the telescope consisted of a Bose–Einstein condensate of two photons, and we then calculated the average number N_0 of photons of energy $E_\gamma / 2$ that would have been emitted for each pair of photons arriving with total energy E_γ. The energy of the emitted bunch was then on average $N_0 E_\gamma / 2$, and the correction factor used was therefore $N_0 / 2$. Because of the change in occupation number, and hence energy, of the bunch of photons during propagation, the precise meaning of the corrected spectrum $(E_\gamma^2 F(E_\gamma))$ is ambiguous, and caution should be applied when interpreting it. Nevertheless, integrating the corrected luminosity $(E_\gamma^2 L(E_\gamma))$ over $\ln(E_\gamma)$ will give the correct total

---

[1] After submission of this paper a preprint by the HEGRA Collaboration appeared (astro-ph/0006092) which renders the possibility of a Bose–Einstein condensate explanation as being rather unlikely.
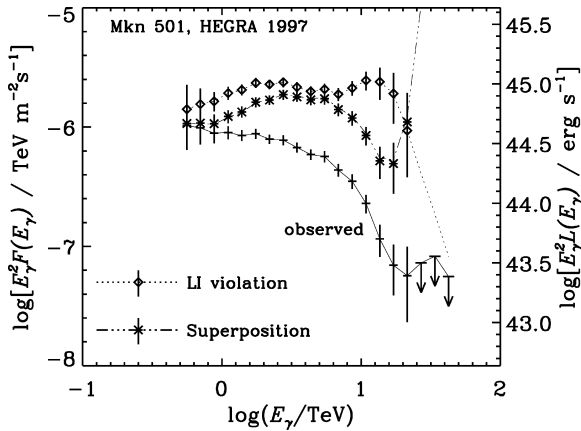
Fig. 4. The time-averaged spectrum of gamma-rays from Markarian 501 observed by HEGRA in 1997 is plotted and compared with the spectrum after correction assuming that each observed "gamma ray" is a coherent superposition of two photons (Superposition), or that Lorentz invariance is violated (LI violation); upper limits have been omitted for clarity from the "corrected" data.

emitted power. Thus, we note that this process is a viable mechanism for solving the problem, and has not yet been ruled out by the observations.

Another possibility which has been suggested is violation of Lorentz invariance [36–39]. Solutions of that kind have also been invoked [40] to explain the observed high energy particles above the GZK [41, 42] cut-off. Amelino-Camelia et al. [37] suggested that the velocity of propagation of energetic particles is modified in vacua due to microscopic quantum fluctuations occurring on scales of the order of the Planck length. Their dispersion relation was used by Kifune [38] to show that intergalactic space will be much more transparent to TeV photons. We follow his treatment, in which energetic photons have velocity $\beta_\gamma c = (1 - \xi_\gamma E_\gamma / E_0)c$, and momentum p given by $p^2 c^2 = (E^2 + \xi_\gamma E_\gamma^3 / E_0)$ from which we can define an effective mass m_γ for high energy photons by $m_\gamma^2 c^4 = -\xi_\gamma E_\gamma^3 / E_0$. Taking $\xi_\gamma = 1$ and $E_0 = 1.2 \times 10^{19}$ GeV (Planck mass) [36,37], and using β_γ and m_γ^2 in Eq. (1), we get the result shown by the thick chain line in Fig. 2. One should be cautious, however, as Eq. (1) was derived assuming s to be Lorentz invariant. Nevertheless, our result gives an indication of the likely effect of Lorentz Invariance violation. Applying our result to correct the HEGRA 1997 Markarian 501 data we obtain the spectrum shown by the diamonds

in Fig. 4, and find that indeed the absorption in intergalactic space is sufficiently reduced to solve the problem, resulting in a spectrum that is largely consistent with both leptonic and hadronic models.

However, it is remarkable that an acceptable source spectrum for Markarian 501 is obtained simply by choosing the Planck mass as the energy scale for violation of Lorentz Invariance. Limits for this energy scale have been obtained previously by looking for time delays in variable sources such as Gamma Ray Bursts [43] and Markarian 421 [44]. By observing instead the energy spectrum of distant sources like Markarian 501 above 10 TeV, we may also have a sensitive tool for exploring the energy scale of Lorentz Invariance violation.

New data on Markarian 501 and Markarian 421 relevant to the issue discussed in this paper are soon to come as both sources are observed to be in a high flaring state in the ongoing observation period (May 2000). Detecting TeV gamma-rays from a more distant BL Lac object (e.g., $z = 0.03$–0.09) would provide for a much needed additional constraint on the phenomenon observed at Markarian 501.

Acknowledgements

References

[1] J.R. Primack, J.S. Bullock, R.S. Somerville, D. MacMinn, Astroparticle Physics 11 (1999) 93.
[2] M.A. Malkan, F.W. Stecker, Astrophys. J. 496 (1998) 13.
[3] E. Dwek et al., Astrophys. J. 508 (1998) 106.
[4] S. Fall, S. Charlot, Y.C. Pei, Astrophys. J. 464 (1996) L43.
[5] D. Elbaz et al., Astron. Astrophys. 351 (1999) L37.
[6] E.L. Wright, astro-ph/0004192;
E.L. Wright, E.D. Reese, astro-ph/9912523.
[7] A. Biviano et al., in: A. Mazure, O. Le Fevre, V. Le Brun (Ed.), Proceedings of the 1999 Marseille IGRAP Conference Clustering at High Redshifts, astro-ph/9910314.
[8] D.P. Finkbeiner, M. Davis, D.J. Schlegel, astro-ph/0004175, submitted to Astrophys. J.
[9] M.G. Hauser et al., Astrophys. J. 508 (1998) 25.
[10] Y. Wang, P.L. Biermann, Astron. Astrophys. 356 (2000) 808.
[11] M. Punch et al., Nature 358 (1992) 477.

[12] F.W. Stecker, O.C. Jager, M.H. Salamon, Astrophys. J. 390 (1992) L49.

[13] R.J. Protheroe, T. Stanev, Mon. Not. R. Astron. Soc. 264 (1993) 191.

[14] R.J. Gould, G. Schreder, Phys. Rev. Lett. 16 (1966) 252.

[15] J.V. Jelley, Phys. Rev. Lett. 16 (1966) 479.

[16] R.J. Protheroe, C.L. Bhat, P. Fleury, E. Lorenz, M. Teshima, T.C. Weekes, in: M.S. Potgieter et al. (Eds.), Proc. 25th International Cosmic Ray Conference, 1997, Durban, Vol. 8, World Scientific, Singapore, 1998, p. 317.

[17] T. Stanev, A. Franceschini, Astrophys. J. 494 (1998) 159.

[18] F.A. Aharonian et al., Astron. Astrophys. 349 (1999) 11.

[19] J.M. Jauch, F. Rohrlich, The Theory of Photons and Electrons, Addison-Wesley, 1955.

[20] W. Bednarek, R.J. Protheroe, Mon. Not. R. Astron. Soc. 310 (1999) 577.

[21] F.A Aharonian et al., Astron. Astrophys. 349 (1999) 29.

[22] F.A. Aharonian et al., Astrophys. J. 539 (2000) 317.

[23] B. Wiebel-Sooth, P.L. Bierman, in: Landolt-Börnstein, Handbook of Physics, Springer, 1998.

[24] F.A. Aharonian et al., Phys. Rev. D 59 (1999) 092003.

[25] M. Catanese et al., Astrophys. J. Lett. 487 (1997) L143; F.W. Samuelson et al., Astrophys. J. Lett. 501 (1998) L17.

[26] A. Djannati-Atai et al., Astron. Astrophys. 350 (1999) 17.

[27] K. Mannheim, S. Westerhoff, H. Meyer, H.-H. Fink, Astron. Astrophys. 315 (1996) 77.

[28] J. Guy, C. Renault, F.A. Aharonian, M. Rivoal, J.-P. Tavernet, Astron. Astrophys. 359 (2000) 419.

[29] D. Petry et al., Astrophys. J. 536 (2000) 742.

[30] R.M. Sambruna et al., Astrophys. J. 538 (2000) 127.

[31] K. Mannheim, Astron. Astrophys. 269 (1993) 67.

[32] A.M. Mücke, R.J. Protheroe, in: GeV–TeV Astrophysics: Toward a Major Atmospheric Cherenkov Telescope VI, Snowbird, Utah, August 1999, in press, astro-ph/9910460; A. Mücke, R.J. Protheroe, astro-ph/0004052.

[33] F.A. Aharonian, astro-ph/0003159.

[34] F.A. Aharonian, P.S. Coppi, H.J. Völk, Astrophys. J. 423 (1994) L5.

[35] M. Harwit, R.J. Protheroe, P.L. Biermann, Astrophys. J. 524 (1999) L91.

[36] S. Coleman, S.L. Glashow, Phys. Rev. D 59 (1999) 116008.

[37] G. Amelino-Camelia, J. Ellis, N.E. Mavromatos, D.V. Nanopoulos, S. Sarkar, Nature 393 (1998) 763; see also: G. Amelino-Camelia, T. Piran, S. Sarkar, hep-ph/0006210.

[38] T. Kifune, Astrophys. J. Lett. 518 (1999) L21.

[39] W. Kluzniak, astro-ph/9905308; R. Aloisio, P. Blasi, P.L. Ghia, A.F. Grillo, Phys. Rev. D 62 (2000) 053010.

[40] H. Sato, astro-ph/0005218.

[41] K. Greisen, Phys. Rev. Lett. 16 (1966) 748.

[42] G.T. Zatsepin, V.A. Kuzmin, JETP Lett. 4 (1966) 78.

[43] B.E. Schaefer, Phys. Rev. Lett. 82 (1999) 4964.

[44] S. Biller et al., Phys. Rev. Lett. 83 (1999) 2108.