Lecture III. Clusters of Galaxies

Guido Chincarini[1]

1. Universita degli Studi di Milano Bicocca Physics Dept. Beppo Occhialini Osservatorio astronomico di Brera – Emerito

This article summarizes the third of a series of lectures delivered at the Cosmology School "Introduction to Cosmology". It reviews the problem of Dark Matter in clusters of galaxies in its historical context.

As I mentioned in my first lecture the missing mass problem started with Zwicky and it has been a rather isolated claim for quite some time. He estimated the mass of clusters of galaxies (Coma cluster) using the velocity dispersion of their member galaxies assuming, due to the simple fact the clusters exist, they are in dynamical equilibrium.

The paper by Humason et al. (1956) gives the status of extragalactic observations about 23 years later. In this first extensive catalogue of Nebulae, a milestone for extragalactic astronomy, 620 objects were observed at Mount Wilson and Palomar and 300 at Lick Observatory. The next step was carried out by Mayall (1960) in his article "Advantages of electronic photography" (CNRS conference). In this work Mayall adds to the existing Coma cluster 18 more redshifts for a total of 50. Part of the project was aimed to search for variations of the velocity dispersion as a function of the distance from the cluster's center to eventually separate galaxies of the clusters from the "field", where galaxies at large distances, intermingle. A real vision! The project was part of a cooperative observational – theoretical effort of Nicholas Mayall with the very well known statisticians of the Statistical Laboratory in Berkeley J. Neyman and Elizabeth Scott so that, in order to allow statistical studies, the sample had to be well defined. For this reason Mayall selected galaxies whose spectra could be obtained with the Crossley nebular spectrograph, in about 2–3 hours. This is a criterion clearly based on the surface brightness of the galaxies and also testify that at the time the concept of a magnitude limited or volume limited sample was not yet viable. The other important point, the limited number of redshifts couldn't lead to reach the goal of the project, Nicholas Mayall made was the need of hundreds of redshifts to fully understand the dynamics of the cluster, however with about 10 redshifts per year the task would be impossible. He specifically refers then to the use of electronics devices, as at the time Merle Walker (he used the Lallemand camera) and Jerry Kron (the Kron camera had the advantage to preserve the photocathode during the change of photographic nuclear plates) were starting at Lick Observatory, as the only way to go.[1] In 1965 Herbert J. Rood got his PhD at the University of Michigan with the thesis: "The dynamics of the Coma cluster of galaxies". I never had a copy of his thesis, I didn't ask at the time[2], and later he published (Rood et al., 1972) a fundamental paper in collaboration with Page, Kintner and King. He told me it took sometime before submitting it since Ivan King

[1]After the talk Charles Fehrenbach (Haute Provence) asked a question (in French!) not in english as he usually did in most meetings. I noticed this since a few years later I participated to a meeting at Haute Provence and again Feherenbach spoke only in French. This remained in my mind as something strange.

[2]I wrote (05/31/2016) to the University of Michigan asking for a copy, I didn't get it.

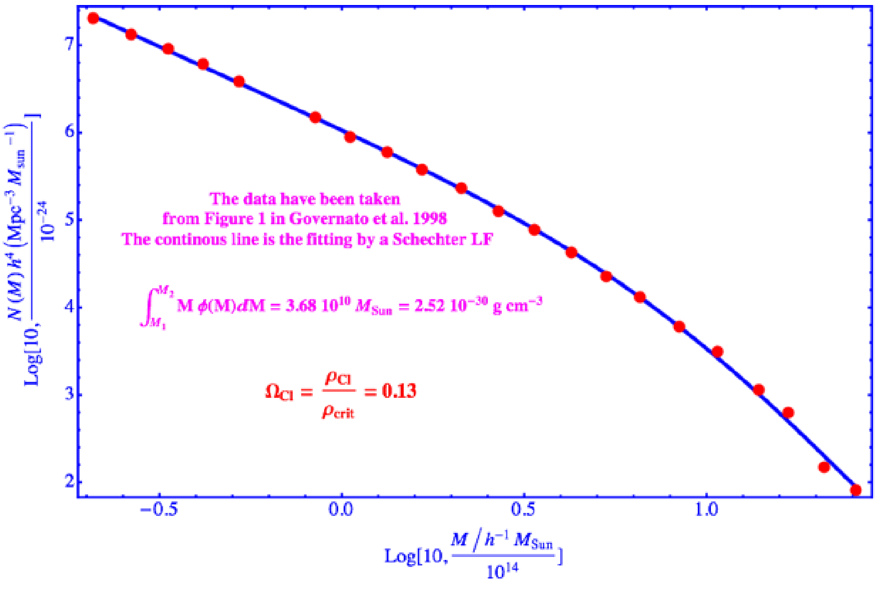

Fig. 1: Cluster mass function and Schechter fit. See Fig. 1 in Governato et al. (1999)

went deeply into it checking all the details. Herbert was terribly interested in this cluster, in the dynamics and the related mass discrepancy (the difference between the optical and dynamical mass), that is the accurate estimate of the Dark Matter mass, and later started, in collaboration with Thornton Page, a program at KPNO to observe galaxies in the Coma cluster. This also because at that time there was not much agreement on what was going on in clusters[3]. At that time, late sixties early seventies, the magnetically focussed RCA image tubes (among others) were used giving the possibility, as foreseen by Mayall, to get spectra of Nebulae (good enough for redshfts) with relatively short exposures. After I got at the Johnson Space Flight Center Houston (I was asked to join the group Thornton Page was forming there)[4] Thornton suggested I join (to substitute him) Herb in the KPNO observations of the Coma cluster. I did that and later after reading the paper by Humason et al. (1956) I noticed that various clusters and in particular the Perseus Pisces cluster, had very few redshifts (three for the Perseus cluster) so that I proposed to Herb to extend the survey work, we discussed and agreed on that matter. I mention all of this, even if of little importance for the science, both to personalize a little bit the lectures and because I due to Herb my getting into the research on Clusters of galaxies. This led us later on into the searching and some pioneering understanding and definition of the Large Scale Structure.[5] We used the Zwicky catalogue Zwicky

[3]See Dekel & Shaham (1979) for a discussion of the high M/L ratio stars in the halo of the spiral NGC 4565 and the unusual interpretation of cluster data by William Tifft.

[4]The plan was to design and eventually build a telescope and relative instrumentation to put on the Moon using an Apollo flight.

[5]In this period the community was very active and, to avoid misunderstanding, here I do not plan a review but simply transfer some personal experience to the students. For details on this matter see, in addition to the many reviews written, also Chincarini (2013).

et al. (1961-1968) and started to list and observe magnitude limited samples[6]. To my knowledge we were among the first doing this. At the time Herb Rood kept most of the connections with the community, I was a new comer and very busy with personal matters, in that period I never participated to a science meeting. Zwicky (1933) based his analysis on the viral theorem and that is what we all did after Zwicky. That was the way to estimate the dynamical mass of Clusters of Galaxies. Later on, and after the detection of X-ray emission, masses will be obtained both from the analysis of the X-ray emission and strong and weak lensing. Clusters, as we learned later on, cover a large range of masses and, as shown by the simulations and work by Governato et al. (1999), the mass function is fit quite well by the analytic Press–Schechter function (Fig. 1).[7]

1 The Virial Theorem from various point of view

We consider a cluster of Galaxies to be a closed system, that is a system of particles (galaxies) interacting with one another but not with other bodies[8]. This is true in a very good approximation in spite of the fact that clusters are part of the large scale structure in which they are embedded. The kinetic energy, a quadratic function of second degree of the velocities, is an homogeneous function and we can write (Euler's theorem) $\sum \bar{\nu_i} \frac{\partial T}{\partial \nu_i} = 2T$ and using $\frac{\partial T}{\partial \bar{\nu_i}} = \bar{p_i}$ we have $2T = \sum\limits_i \bar{p_i}\bar{\nu_i} = \frac{d}{dt}\left(\sum\limits_i \bar{p_i}\bar{r_i}\right) - \sum\limits_i \bar{r_i}\bar{p_i}$. The physical system we are considering is located in a definite region of space where the motion of the galaxies occur with finite velocities. It is a closed system where the range of parameters is finite and therefore characterized by bounded functions. If $F(t)$ is a bounded function, then the mean of the derivative is zero, that is:

$$f = \lim_{\tau\to\infty} \frac{1}{\tau} \int\limits_0^\tau \frac{dF}{dt} dt = \lim_{\tau\to\infty} \frac{F(\tau) - F(0)}{\tau} = 0.$$

The function $\frac{d}{dt}(\sum\limits_i \bar{p_i}\bar{r_i})$ is the derivative of a bounded function so that its mean value is zero. Furthermore according to Newton's law $\bar{p} = -\frac{\partial\Omega}{\partial\bar{r}}$ (Ω is the potential energy) and we have $2\langle T\rangle = \langle\sum\limits_i \bar{r_i}\frac{\partial\Omega}{\partial\bar{r_i}}\rangle$. If the potential energy is a homogeneous function of the coordinates, that is if $\Omega(\alpha\bar{r_1}, \alpha\bar{r_2}, \ldots, \alpha\bar{r_n}) = \alpha^k\Omega(\bar{r_1}, \bar{r_2}, \ldots, \bar{r_n})$, where k is

[6]The magnitudes in the Zwicky's catalogues were rather accurate (better than 0.5 mag). Few mistakes even if some astronomers seemed to complain with him about that. Once, while we were having lunch at Caltech when I visited him, he remarked that the ones who complain should try to measure millions of galaxies and check if they make any mistake! At the time we didn't have yet digitised plates.

[7]The red dots have been obtained by measuring the points in the figure by Governato et al. (1999)

[8]Not completely true. We know nowadays that clusters are kind of knots in the Large intermingled filamentary structure. Zwicky believed they were superposed to an almost uniform background of galaxies (his related point was that superclusters do not exist since he never saw a cluster of clusters similar to what we see for a cluster of galaxies). On the other hand see the cluster agglomerate discovered by Scaramella et al. (1989) and later improperly named the Shapley concentration, Raychaudhury (1989).

called the degree of homogeneity, using Euler's theorem $\left[\sum_i^n \frac{\partial f}{\partial x_i} x_i = kf(x_i)\right]$ we have $2\langle T\rangle = k\langle\Omega\rangle$. For a Newtonian interaction the potential energy is proportional to $\mathbf{r}^{-1}$ so that $k=-1$ and $\langle 2T\rangle = -\langle\Omega\rangle$ and $E = \langle T\rangle + \langle\Omega\rangle = \langle T\rangle - \langle 2T\rangle = -\langle T\rangle$ with these mean values intended over time[9]. Discussing the estimates of the mass in the solar neighbourhood we treated the stars as a system of particles subjected to a smooth potential. The equation governing such systems are the collisionless Boltzmann and Jeans equations where the estimate of the distribution function (averaged in this case over time) would give full knowledge of the system. The derived tensor virial equation is valid also for a system of N mutually gravitating particles:

$$\frac{1}{2}\frac{d^2}{dt^2} I_{jk} = 2K_{jk} + W_{jk}.$$

From which, assuming the system is in a steady state or is changing very slowly on the time scale of interest, $\frac{d^2}{dt^2} I_{jk} = 0$, we can derive the Virial theorem in the form given above. We discuss now briefly the velocity dispersion in a spherical system (the geometry we believed a cluster has) because of some analogies with the derivation of the mass using the X-ray observations. The Jeans equation for a spherically symmetric system in a steady state, $\langle\nu_{\rm r}\rangle = \langle\nu_\theta\rangle = 0$, is $\frac{d}{dr}(\rho\langle\nu_{\rm r}^2\rangle) + \frac{\rho}{r}\left[2\langle\nu_{\rm r}^2\rangle - (\langle\nu_\theta^2\rangle + \langle\nu_\varphi^2\rangle)\right] = -\rho\frac{d\Phi}{dr}$. There is no evidence (except for some indications in a few cases) that clusters of galaxies are flattened and rotating so that we can write $\langle\nu_\theta^2\rangle = \langle\nu_\varphi^2\rangle$ and $\frac{1}{\langle\nu_{\rm r}^2\rangle}\frac{d}{dr}\left(\rho\langle\nu_{\rm r}^2\rangle\right) + 2\frac{\rho}{r}\left[1 - \frac{\langle\nu_\theta^2\rangle}{\langle\nu_{\rm r}^2\rangle}\right] = -\frac{\rho}{\langle\nu_{\rm r}^2\rangle}\frac{d\Phi}{dr}$. Defining a degree of anisotropy as $\beta \equiv 1 - \frac{\langle\nu_\theta^2\rangle}{\langle\nu_{\rm r}^2\rangle}$ with a potential $\frac{d\Phi}{dr} = \frac{GM(r)}{r^2}$ we derive for the mass $M(r) = -\frac{r\langle\nu_{\rm r}^2\rangle}{G}\left(2\beta + \frac{d\ \ln\rho}{d\ \ln r} + \frac{d\ \ln\langle\nu_{\rm r}^2\rangle}{d\ \ln r}\right)$. Observationally we can only get two of the three parameters involved so that we are constrained to a model and in any case the estimate of the anisotropy parameter is difficult and not motivated by the observations. A spherical ellipsoid of velocities is the obvious solution and this implies $\beta = 0$. The derived equation is equal to what we would get for a gas in hydrostatic equilibrium with the temperature substituting the velocity dispersion.

We clearly measure the projection of the luminous density on the plane of the sky and the radial velocity dispersion that, thanks to the assumption of isotropy, is the same on the three axes (see Fig. 2). The projected luminosity density, this is the observable, and matter is the column density, that is

$$I(R) = 2\int_0^\infty \rho(z)dz = 2\int_0^\infty \rho(\sqrt{r^2-R^2})\frac{r\,dr}{\sqrt{r^2-R^2}}$$

and for the velocity dispersion

$$I(R)\sigma^2(R) = 2\int_0^\infty \rho(\sqrt{r^2-R^2})\frac{r\langle\nu_{\rm r}^2\rangle dr}{\sqrt{r^2-R^2}}\,.$$

These are Abel integral equations from which we can derive $\rho(r)$ and $\rho(r)\langle\nu_{\rm r}^2\rangle$.

[9]Landau & Lifshitz (1976). For the derivation of the tensor virial theorem is given in all details by Binney & Tremaine (1987).

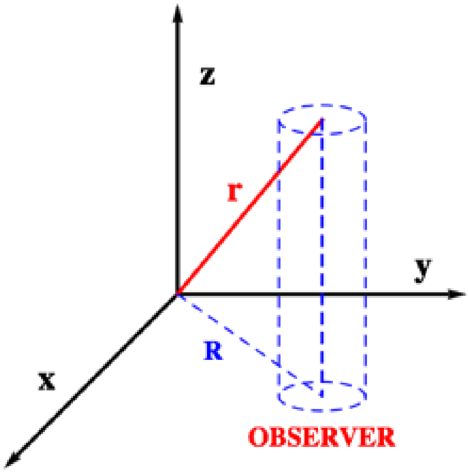

Fig. 2: Column density and related coordinates.

An other simple and rather illuminating way to derive the viral theorem is, following Eddington (1916), as done in Ogorodnikov (1965). The moment of inertia of a system of massive objects respect their center of mass is:

$$I = \sum_{i=1}^{N} m_i r_i^2 = \sum_{i=1}^{N} m_i (x_i^2 + y_i^2 + z_i^2)\,,$$

and assuming a steady state $\frac{d^2}{dt^2} I = 0$ the moment of inertia either remains constant or change at a constant rate. Assuming it changes slowly expanding in power of time and keeping the first term, $I \approx a + bt$ a linearly non steady system.

By differentiating the above equation we have

$$\frac{1}{2}\frac{d^2}{dt^2} I = \sum_{i=1}^{N} m_i(\dot{x}_i^2 + \dot{y}_i^2 + \dot{z}_i^2) + \sum_{i=1}^{N} m_i(x_i \ddot{x}_i + y_i \ddot{y}_i + z_i \ddot{z}_i) = 0\,.$$

Indicating the potential energy of the system as $\Omega = -G \sum\limits_{i=1}^{N} \frac{m_i m_j}{r_{ij}}$ where $r_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2}$ from Newton equations $m_i \ddot{x}_i = -\frac{\partial \Omega}{\partial x_i}$, $m_i \ddot{y}_i = -\frac{\partial \Omega}{\partial y_i}$, $m_i \ddot{z}_i = -\frac{\partial \Omega}{\partial z_i}$ we have $\sum\limits_{i=1}^{N} m_i(x_i \ddot{x}_i + y_i \ddot{y}_i + z_i \ddot{z}_i) = -\sum\limits_{i=1}^{N} \left(x_i \frac{\partial \Omega}{\partial x_i} + y_i \frac{\partial \Omega}{\partial y_i} + z_i \frac{\partial \Omega}{\partial z_i} \right)$. From Euler's Theorem $-\sum\limits_{i=1}^{N} \left(x_i \frac{\partial \Omega}{\partial x_i} + y_i \frac{\partial \Omega}{\partial y_i} + z_i \frac{\partial \Omega}{\partial z_i} \right) = -\Omega$ and since $\sum\limits_{i=1}^{N} m_i(\dot{x}_i^2 + \dot{y}_i^2 + \dot{z}_i^2) = 2T$ we have $2T + \Omega = 0$. The application Rood and I did in the early work on clusters of galaxies, see for instance Chincarini et al. (1975), was to determine the mass using the following relations: $\sum m_i \nu_i^2 = -G \sum \frac{m_i m_k}{r_{ik}}$ with $M = \sum m_i$, $\frac{1}{M} \sum \frac{m_i}{M} \nu_i^2 = -G \sum \frac{m_i}{M} \frac{m_k}{M} \frac{1}{r_{ik}}$ and $M = \frac{V_{\mathrm{D}}^2 R}{G}$ where V_{D} and R are defined as $V_{\mathrm{D}}^2 = \sum \frac{m_i}{M} \nu_i^2$ and $R = \sum \frac{M^2 r_{ij}}{m_i m_k}$. If the ν_i refer to the observed velocities and the separations are measured as projected on the sky, we have, accounting for the projection (Limber, 1960) and tree degrees of freedom:

$$V_{\mathrm{D}}^2 = 3\left\{\frac{\sum m_i (V_i - V_{\mathrm{cm}})^2}{M} - \frac{m_i \sigma^2(V_i)}{M}\right\},$$

$$V_{\mathrm{cm}} = \frac{\sum m_i V_i}{M},$$

$$R = \frac{\pi \frac{M^2}{2}}{\frac{1}{2}\sum\sum \frac{m_i m_j}{r_{ij}}},$$

$$M = \sum m_i = \sum L_i f_i .$$

Generally the luminosities are available and not the masses, for this reason the weights are derived using the mass to luminosity ratio for the different morphological types. For the analysis errors see also Danese et al. (1980).

2 The cluster mass

As mentioned earlier the paper on the Coma cluster by Rood et al. (1972) is a milestone and gives a detailed account of the velocity dispersion following the work by Mayall (1960). In that paper two important points are made: 1) "Thus we are forced to conclude that the gravitational field that holds the Coma cluster together is provided by an unseen mass about 7 times as large as that of the galaxies that we see." and 2) "Thus the missing mass is present but is not in the galaxies. We can show, however, that its distribution in the cluster is very similar to that of the galaxies." After the begging of the surveys by Rood and by Chincarini and Rood, the amount of data collected in the field increased tremendously.

Colless & Dunn (1996) carried out the analysis of the region of the Coma cluster using 552 redshifts. These authors measure a mass of $0.9 \times 10^{15} h^{-1} M_\odot$ in excellent agreement[10] with a previous X-ray measurement (ROSAT) made by Briel et al. (1992) and also find, as detected in X-ray, substructures, (Fig. 3).

Clearly with the rapid progress of the technology the number of observed redshifts grew exponentially so that it had been possible to tackle with high accuracy more ambitious projects: see for instance Geller et al. (1999) and Łokas & Mamon (2003). The latter used data from the literature.

A large sample (Canadian Network for Observational Cosmology – CNOC) came about in 1996, 1997, by Carlberg et al. (1996, 1997) who based their analysis on 2600 velocities in 16 high X-ray luminosity clusters.

$$\left\langle \frac{M}{L} \right\rangle_{\mathrm{16Clusters,vir}} = 289 \pm 50 h \frac{M_\odot}{L_\odot},$$

$$\left\langle \frac{M}{L} \right\rangle_{\mathrm{field}} = 215 \pm 59 h \frac{M_\odot}{L_\odot},$$

$$\Omega_0 = 0.24 \pm 0.05 \pm 0.09 .$$

[10] Note that Briel et al. (1992) use h_{50}^{-1} while Colless & Dunn (1996) use the correct definition ($H_0/100 = h$).

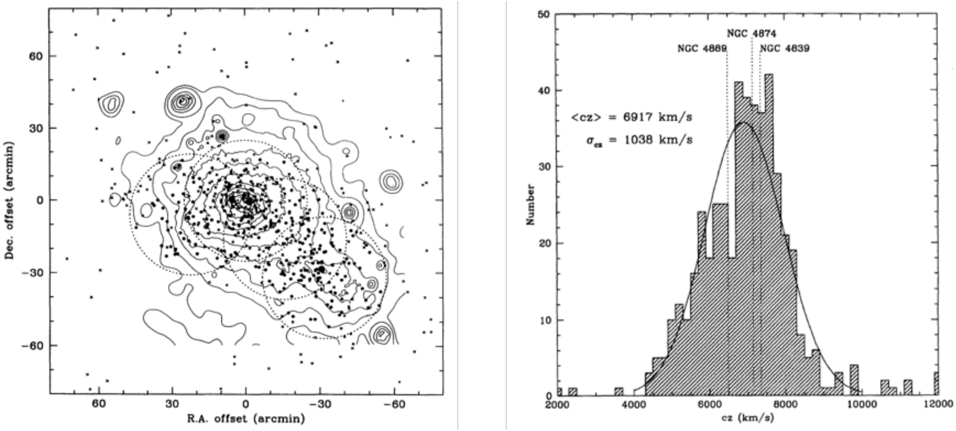

Fig. 3: Left: The distribution of galaxies in the Coma Cluster. The contours represent density levels of the luminosity distribution in the soft X-ray, Right: Distribution of redshifts in the region of the Coma cluster. The analysis of the redshift distribution also evidences two sub-clusters: the main cluster $< cz >= 6853\ \mathrm{km\,s^{-1}}$ and $\sigma_{\mathrm{cz}} = 1082\ \mathrm{km\,s^{-1}}$ and the subgroup centered on NGC 4839 with $< cz >= 7339\ \mathrm{km\,s^{-1}}$ and $\sigma_{\mathrm{cz}} = 329\ \mathrm{km\,s^{-1}}$ (Colless & Dunn, 1996).

Their result, very close to what we find today, gives, out to a radius where the density is about 200 the critical density $\langle\frac{M}{L}\rangle = 295(289 \text{ in } 1997) \pm 53(50) h \frac{M_\odot}{L_\odot}$. The reason to limit to the radius where $\rho \sim 200\rho_{\mathrm{c}}(z)$, with $\rho_{\mathrm{c}}(z) = \rho_{\mathrm{c}}(0)(1+z)^3$, is mainly due to the fact that at larger radii the dispersion is sensitive to infall velocities and that is why this radius is generally referred to as the viral radius. These values are in excellent agreement with the estimates of mass via X-ray emissions by Hughes (1989) who estimated for the Coma cluster $M_{\mathrm{vir}}(< 5h_{50}^{-1}\mathrm{Mpc}) = 1.85 \pm 0.24 \times 10^{15} h_{50}^{-1} M_\odot$ and $\frac{M}{L} = 165 \pm 25 h_{50} \frac{M_\odot}{L_\odot}$[11]. Similar results were obtained earlier by Kent & Gunn (1982). These authors develop a rather detailed analysis building on the early work by Rood et al. (1972) and by Mayall (1960) and fit quite well the distribution of galaxies using a King model (1966) with a core radius of $340 - 400$ kpc. Within 7.2 Mpc they estimate $M_{\mathrm{vir}}(< 7.2h_{50}^{-1}\mathrm{Mpc}) = 2.9 \times 10^{15} h_{50}^{-1}\ M_\odot$ and $\frac{M}{L} = 181$. The analysis rules out the existence of supermassive objects $M \geq 10^{14} M_\odot$ in the centre and most important they insist on the need of measuring redshifts at large distances from the cluster centre to improve, among other things, the measure of the velocity dispersion as a function of the distance from the centre. The velocity dispersion decreases at large distances and the distribution of Dark Matter likely follows the distribution of galaxies.

Merritt (1987) analysis of various models within 1 Mpc from the centre of the Coma cluster is consistent with a mass of 6×10^{14} Mpc ($H_0 = 100$ km s^{-1} Mpc^{-1}), $\frac{M}{L} = 350$, and however there does not seem to be any strong evidence from the kinematical data alone that the dark and luminous components of the Coma cluster

[11]It is unfortunate that after it was agreed (IAU) to use h in units of 100 ($h = H_0/100$) to avoid the confusion due to the use of different H_0 by the various authors, many complicated further the field (or the confusion) by giving h in units of 50 and others in units of 75.

are similarly distributed. Łokas & Mamon (2003) approaches the mass distribution in the Coma cluster in a very interesting way using the distribution of galaxies and gas by deriving the distribution of Dark Matter using the velocity moments derived from the sample of velocities of $355 E - SO$ galaxies. Lokas gives details of the analytical fits she obtained and reproduces the observed velocity moments with mass given by the various components of matter: Galaxies, Gas and Dark Matter, she derives the running of the various components of the mass as a function of the distance from the cluster centre. Not only we have agreements with estimates in the X-ray and weak lensing but the density parameter is very close to what is accepted nowadays.

In 1966 we have the first observations on clusters of galaxies in the X-ray band. Boldt et al. (1966) using a balloon flight launched from the Holloman Air Force base in New Mexico detected a source of X-ray (very likely the Coma cluster of galaxies) at about 25 keV with a photon number of $\sim 10^{-2}\mathrm{cm}^{-2}\mathrm{s}^{-1}\mathrm{keV}^{-1}$. Felten et al. (1966) discusses the results proposing a thermal bremsstrahlung model. On December 12, 1970 (Kenya's Independence Day), the group of Riccardo Giacconi at the AS&E with Herbert Gursky (read his paper in Exploring the Universe: A Festschrift in honour of Riccardo Giacconi) and Bruno Rossi launched the Uhuru (freedom in the Swahili dialect) satellite from the Italian platform San Marco that is located off the coast of Kenya[12]. Soon after the launch the evidence that the clusters had a large amount of intergalactic gas became very strong. Important observations were gained also at radio wavelengths. Miley et al. (1972) discussed the radio tails galaxies and proposed they represent radio trails of active galaxies along trajectories through dense intergalactic medium. Later Jaffe & Perola (1973) developed a detailed dynamical model[13]. These radio observations and analysis of tail galaxies in clusters of galaxies are the first robust evidence of the existence of a intergalactic gas in clusters. Indeed reporting Miley ... "The component distortion, the displacement of the optical galaxy to the front edge of the radio brightness distribution and the direction and form of the radio trail all suggest motion of a radio galaxy resisted through a gaseous medium" and ... "For galaxy speeds as high as 2500 $\mathrm{km\,s^{-1}}$ the above inequality and these internal pressures lead to lower limits of intergalactic gas densities of $1.6\times10^{-28}\ \mathrm{g\,cm^{-3}} - 2.2\times10^{-27}\ \mathrm{g\,cm^{-3}}$". The physics and the model of the ram pressure is rather well understood, nevertheless for clarity we reproduce here (but see also the solar wind bow shock and tail) the sketch made by Jaffe & Perola (1973) – Fig. 4.

The direct evidence came with the detection of the X-ray emission. Gursky et al. (1972) using the Uluru data characterised the gas as due to thermal bremsstrahlung at a temperature of about 7.3×10^7 K and a mass of 3×10^{13} solar masses, about 1% of the mass needed to stabilise the cluster. The stage had been set: a) the intergalactic gas emits thermal bremsstrahlung, b) the mass of the gas however is not enough to account for the dynamical mass and c) X-ray data will become an important tool to estimate cluster's physical parameters.

[12] I visited (2001) the platform shortly before the launch of the Swift Satellite when we decided, after I proposed it to ASI, to use the Malindi ASI Ground Station for the Swift Satellite. See presentation.

[13] NGC 1265 due to its high random velocity component was a very clear indication, I was told, for the Miley et al. model. Since I could not get a second spectrum to make sure of the redshift, I asked Vera Rubin who had telescope time, to check. She did and she phoned me everything was OK. Thanks Vera.

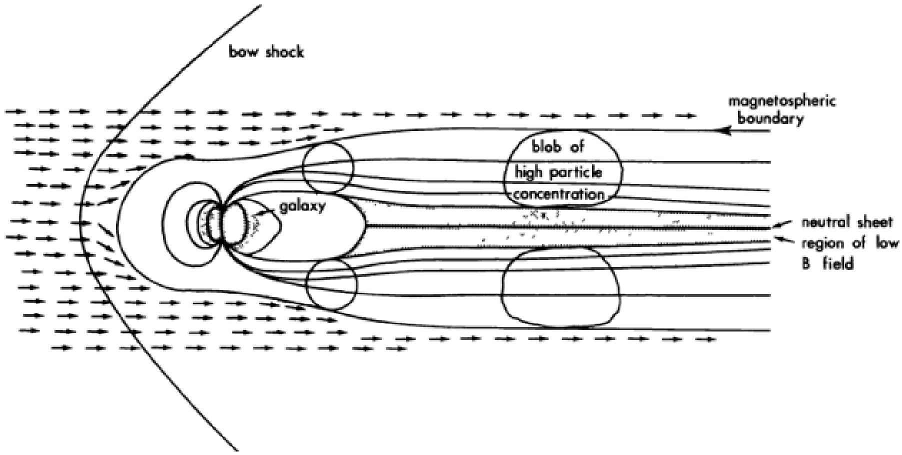

Fig. 4: Sketch of the magnetospheric model (their model II). The various structures caused by the interaction of the galactic magnetic field, the intergalactic gas, and the concentrations of relativistic particles released by the galaxy. A strong indirect evidence for the existence of intergalactic gas in clusters of galaxies.

Soon after the detection of the X-ray gas in clusters it was uncertain whether the emission mechanism was inverse Compton or Bremsstrahlung. It is in this period that Sunyaev & Zeldovich (1972) proposed the theory of what nowadays is called the SZ effect as a test for identifying the emission mechanism of the gas. The idea is of the simplicity that characterise the great intuitions.

If a mass of hot gas is embedded in the MWB and located between us and the recombination era then due to the Thomson scattering some of the MWB photons will be deviated from their path by the electrons of the gas and we will see a "hole" on the MWB at the location of the cluster. The gas distorts the relic radiation spectrum. What a vision!

The article (Comments on Astrophysics and Space Physics, Vol. 4, p.173) give all the needed details and Fig. 5 reproduces the sketch made by Sunayev and Zeldvich, simple and very clear.

The gas must be in equilibrium with the mass since otherwise it would either collapse toward the center if too cold (and eventually form galaxies and stars) or if too hot simply escape from the cluster potential (see also Turnrose & Rood, 1970). The order of temperature is easily found under this equilibrium hypothesis assuming the potential of the cluster is the one holding both gas and galaxies together. With a velocity dispersion of the order of $1000\,\mathrm{km\,s^{-1}}$ we have, as an order of magnitude $T = \frac{m_{\mathrm{p}}\sigma^2}{3k} \approx 10^7 - 10^8$ K. In this range of temperatures the main mechanism for the cooling of the gas is thermal bremsstrahlung $\varepsilon^{\mathrm{ff}} = 1.44 \times 10^{-27}\,\overline{g}\,T_{\mathrm{g}}^{1/2} n_{\mathrm{e}} \sum_i Z_i^2 n_i$ erg cm^{-3} s^{-1} using a solar abundance (stopping at the Nickel for the exercise) $\sum Z_i n_i = 1.17\,n_{\mathrm{e}}$, so that we can write $\varepsilon^{\mathrm{ff}} = 1.97 \times 10^{-27}\,T_{\mathrm{g}}^{1/2} n_{\mathrm{p}}^2$ erg cm^{-3} s^{-1}. We estimate an approximated cooling time by simply dividing the energy of the gas by the bremsstrahlung emission $t_{\mathrm{cooling}} = \frac{3knT_{\mathrm{g}}}{1.97\times 10^{-27} T_{\mathrm{g}}^{1/2} n_{\mathrm{p}}^2}$

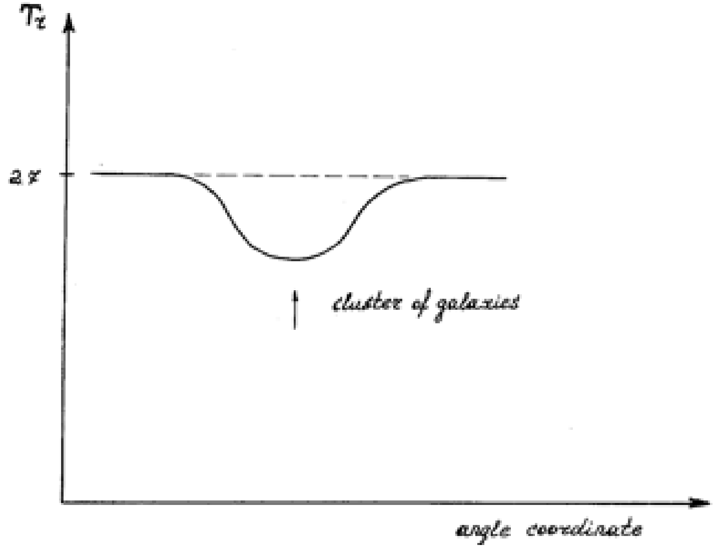

Fig. 5: The "hole" in the microwave background.

$= 6.67 \times 10^3\, T_{\rm g}^{1/2} n_{\rm p}^{-1}$ years; with densities of about 10^{-3} cm^{-3} and temperatures of about 10^8 K the cooling time is of the order of 10^{10} years, so that the cooling time is about larger than the age of the Universe and therefore irrelevant for our purposes. We can use the gas to estimate the mass.

The gas is kept in equilibrium by the gravitational potential and we can write:

$$\frac{1}{\rho_{\rm gas}} \frac{dP}{dr} = -\frac{\partial \Phi}{\partial r} = -\frac{GM}{r^2},$$

where

$$P = kT(r) \frac{\rho(r)}{\mu m_{\rm p}},$$

from this

$$\frac{1}{\rho} \frac{1}{\mu m} \left[\rho \frac{\partial T}{\partial r} + T \frac{\partial \rho}{\partial r} \right] = -\frac{GM}{r^2},$$

$$M(r) = -\frac{kT_{\rm g} r}{G \mu m_{\rm p}} \left[\frac{d \ln \rho_{\rm g}}{d \ln r} + \frac{d \ln T_{\rm g}}{d \ln r} \right].$$

The temperature is measured from the spectrum. The measure of the gradient temperature would require space resolution and a reasonably good flux to measure the temperature at various distances from the cluster center. Most cluster of galaxies are stable systems and, as we have seen, are in first approximation spherically symmetric. Zwicky (1957) in his fascinating "Morphological Astronomy" gives an exhaustive description of the distribution of galaxies in clusters and shows that after normalisation they all fit the distribution of the Lane Emden gravitational isothermal gas sphere. Early counts of galaxies in the Coma cluster have been done by Omer et al. (1965). A perused model of the distribution is that of the isothermal approximation by King (1962) since this allow easy analytical computations. Fitting for a large number of counts of galaxies in clusters has been done by Chincarini (1978). Cavaliere & Fusco-Femiano (1976, 1978, 1981) discussed a physically consistent hydrostatic model for the X-ray observations accounting both for the galaxies and the gas (see presentation – β model). In 1996 Navarro Frenk and White (NFW) (Navarro et al., 1996) published their results of N body simulations where they investigated the structure of dark halos in a cold dark matter cosmogony. In this work they derived a formula that has been perused since then. Recently various empirical fitting formulae as refinements have been published and also some modifications too the NFW profile to be better in agreement with new N-body simulations. Some of these are useful because they refer to observational parameters, others may call to memory the first line of the Morphological Astronomy by Zwicky: "Man has a great tendency to get lost or to hide, as the case may be, in a jungle of details and in unnecessary complications".

3 The light and matter distribution

The isothermal sphere, equation of state $P = k\rho$, is a particular case of the Emden's equation derived for the polytropic spheres. For a complete analysis see Chandrasekhar (1939) (An introduction to the theory of stellar structure, 1939), Ogorodnikov (1965) (Dynamics of stellar systems), Binney & Tremaine (1987) (Galactic Dynamics). Using the equation of the hydrostatic equilibrium:

$$dP = -\rho d\phi = \rho(r)g(r)dr \;\;\&\;\; g(r) = \frac{G}{r^2}\int\limits_0^r 4\pi r^2 \rho(r)dr,$$

$$dP + \rho(r)dr\frac{G}{r^2}\int\limits_0^r 4\pi r^2\rho(r)dr = 0,$$

$$G\int\limits_0^r 4\pi r^2\rho(r)dr + \frac{dP}{dr}\frac{r^2}{\rho(r)},$$

$$G4\pi r^2\rho(r) + \frac{d}{dr}\left(\frac{dP}{dr}\frac{r^2}{\rho(r)}\right) = 0,$$

assuming we are dealing with a gas sphere that obey the equation of state $P = \frac{\rho kT}{\mu}$

$$4\pi G r^2\rho(r) = -\frac{d}{dr}\left(\frac{d\left(\frac{\rho kT}{\mu}\right)}{dr}\frac{r^2}{\rho(r)}\right),$$

$$\frac{d}{dr}\left(r^2\frac{d\ln\rho}{dr}\right) + \frac{4\pi G\mu}{kT}r^2\rho(r),$$

and using dimensionless variables $\rho = \rho_0 e^{\Psi}$, $r = r_0\xi$ $\frac{4\pi G\mu}{kT}r_0^2, \rho_0 = 1$ we derive

$$\frac{d}{d\xi}\left(\xi^2\frac{d\Psi}{d\xi}\right) + \xi^2 e^{\Psi} = 0 \quad \text{or}$$

$$\frac{d^2\Psi}{d\xi^2} + \frac{2}{\xi}\frac{d\Psi}{d\xi} = e^{\Psi}.$$

This equation (Lane–Emden) can be integrated only numerically. The numerical integration and fitting of the data is rather simple. In any case Zwicky (1957) tabulates the density distribution of the reduced (in dimensionless quantities) Isothermal Gas Sphere and of its projection into a plane. Applications and examples can be found in Chincarini et al. (1979).

King (1972), following the study of the Coma cluster in collaboration with Rood et al. (1972) and his work on the density distribution of a family of self gravitating dynamical models, proposes the approximation $\rho(r) = \rho_0 \left(1 + \frac{r^2}{R_{\rm c}^2}\right)^{-3/2}$ where $R_{\rm c}$ is the core radius (3 times the isothermal structural length). This fits quite well the isothermal sphere (and the data) not too far from the central region. A simple integration gives the projected density (this is valid also for the distribution of light) and it can be easily shown that the core radius is the distance from the centre at which the projected density is 50% of the central density. Finally useful for analytical work is also the rather

$$\frac{\left(1 + \frac{x^2}{3^{1/a}12e}\right)^a}{\left(1 + \frac{x^2}{6}\right)^{1+a}\left(1 + (2^{-a} - 1)\left(\frac{bx^2}{1+bx^2}\right)\right)} \exp\left[\frac{a\left(\frac{x^2}{6}\right)}{\left(1 + \frac{x^2}{6}\right)}\right],$$

accurate approximation given by Liu (1996) with $a = 0.551$ and $b = 3.8410^{-4}$. Assuming the gas and the galaxies are in a steady state we can write $\frac{3kT}{\mu} = \sigma^2$ and derive the following relation $R_{\rm c} = \sqrt{\frac{3\sigma^2}{4\pi G\rho_0}}$ that relates the three observables $R_{\rm c}$ (core radius), σ^2 (velocity dispersion or T) and ρ_0 (central density).

The Navarro, Frenk and White (NFW) profile (1996) Navarro et al. (1996) $\rho(r) = \frac{\langle\rho\rangle_{{\rm r}<200} f(c)}{\left(\frac{rc}{r_{200}}\right)\left(1+\frac{rc}{r_{200}}\right)^2}$ where $\langle\rho\rangle_{{\rm r}<200}$ is the mean density within the viral radius r_{200}, c a dimensionless parameter and $f(c)$ a function of c (see the original paper for details). The profiles in the Fig. 6 have been derived using the following parameters: Mass= $1.88 \times 10^{15} M_\odot$ (for Coma, Kubo et al. (2007)) and for ISO $R_{\rm c} = 9'$(245 kpc at $cz = 6539$), $\sigma = 1900$, $a = 0.551$, $b = 3.8410^{-4}$. The profiles differ in the central region. For the baryonic matter the local physics may be determinant in the characterisation of the profile. The NFW profile has been derived from CDM numerical simulations for testing DM, the baryonic matter should follow the DM potential except that shocks and phenomena related to energy dissipation may often dominate the central region of clusters. Often the core of a cluster is dominated by a cD galaxy. Hughes (1989) finds that the optical and X-ray data (Einstein Observatory – Giacconi et al. (1979)) are completely consistent with a model in which the virial mass follows the distribution of visible light; indeed this is the preferred solution he derives. The total virial mass within $5h_{50}^{-1}$ Mpc is $(1.85 \pm 0.24) \times 10^{15} h_{50}^{-1} M_\odot$, and the mass-to-light ratio is $165 \pm 15h_{50}$ in solar units. Briel et al. (1992) use ROSAT data and estimate within $5h_{50}^{-1}$ Mpc a mass $(1.8 \pm 0.6) \times 10^{15} h_{50}^{-1} M_\odot$, the fraction of cluster mass contained in hot gas $0.30 \pm 0.14h_{50}^{-3/2}$ (see Fig. 7).

Lensing is another way to go, the observations obtained in the last years are magnificent and very stimulating. The distortions due to the gravitational field allow to model accurately (see also lecture II) the mass distribution generating them. The

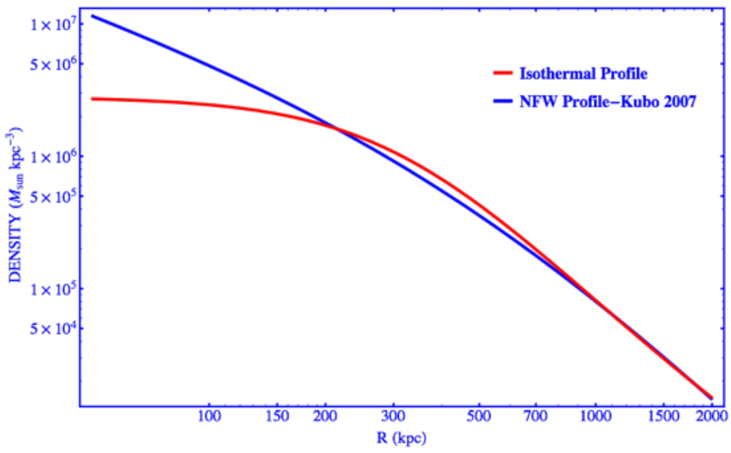

Fig. 6: Comparison between the isothermal profile and the NFW.

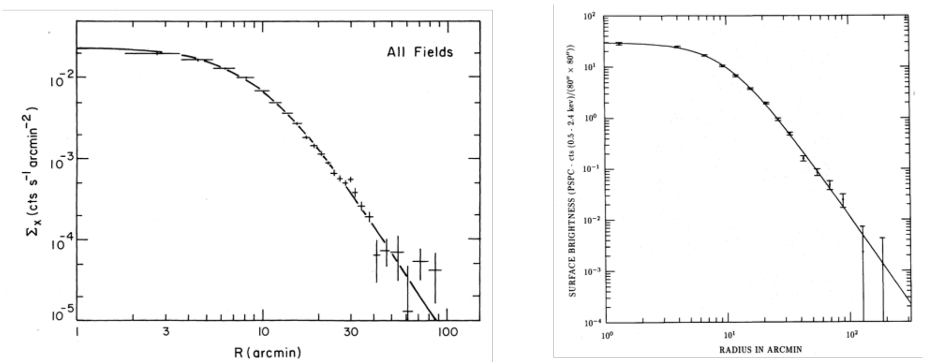

Fig. 7: On the left Einstein Observatory data (Hughes, 1989) and on the right ROSAT data (Briel et al., 1992).

development of this type of observations and related data analysis in the last 20 years has been tremendous due to a) the awareness of the potentiality of the technique, b) the high resolution of telescopes and detectors that kept improving and finally c) the capability of computers and software to handle large amounts of data. The work by Sharon et al. (2012), Jauzac (2016), Kelly (2015) and Goobar (2017) among many others give a feeling for the advanced state of the art in this field. The progress made, the accuracy reached and the possibilities that are at the horizon with JSWT and ALMA are far beyond what we could imagine few years ago. Kubo et al. (2007) uses weak lensing to estimate the mass profile up to a distance $r_{200} = 1.99^{+0.22}_{-0.21}$ Mpc and $M_{200} = 1.88 \times 10^{15} h^{-1} M_{\odot}$. The cluster's density fits quite well the NFW profile. Hoekstra (2007) derive WL masses that are essentially model independent and find that they agree well with measurements of the velocity dispersion of cluster galaxies and with the results of X-ray studies. The mass estimates and the density profiles of

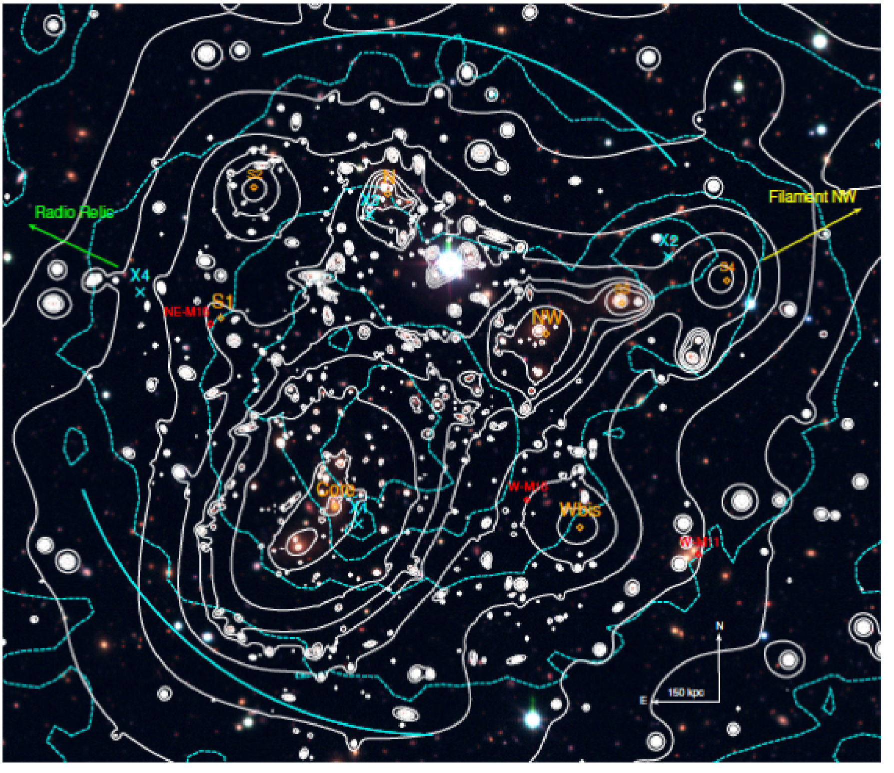

Fig. 8: Contours: White = strong-weak lensing, Cyan=Chandra. Diamonds: Orange=substructures by lensing, red = W & NE clumps, cyan = remnant cores detected in Chandra & Newton XMM. From Jauzac (2016).

the various components become more and more precise as to allow to estimate the detail of the profile and clumping of matter as we can see in Fig. 8 (see also images in the presentation).

Sunyaev & Zeldovich (1972) in the paper we mentioned earlier, states:

> "If hot intergalactic gas really exists in clusters of galaxies, there are all the conditions for action of the proposed effect. Thus, for example, the X-ray radiation of the Coma cluster of galaxies is interpreted as the bremsstrahlung radiation of a hot intergalactic gas having $T_{\rm e} \sim 7 \times 10^7$ K and the density $N_{\rm e} \sim 10^{-3}$ cm^{-3}. The linear dimension of the source is estimated to be. Multiplying these figures, we find
>
> $$\frac{\Delta T_{\rm r}}{T_{\rm r}} = -2y = -2\sigma_{\rm T} N_{\rm e} l \frac{kT_{\rm e}}{m_{\rm e}c^2} \sim 2 \times 10^{-4},$$
>
> the value accessible for observations. Namely, such effect was recently discovered in Coma by Pariysky. The deficit of brightness (hole) in the Coma is difficult to understand by any other mechanism!"

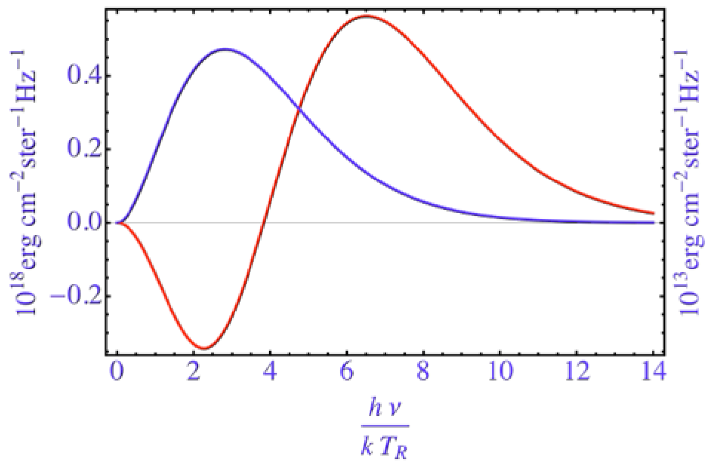

Fig. 9: The blue line is the Black Body at $T_{\rm R} = 2.7$ K. The relevant label is to the right of the frame. The red line represents the SZ spectrum distortion and the related label is on the left of the frame. The SZ effect (red) scale has been multiplied by a factor 10^5 compared to the BB−2.7 K (blue) spectrum.

A very simple and brilliant idea. If I have a clump of gas between us and the MWB we see a hole at the location of the clamp because the background photons interact with the high energy electron of the gas and are shifted to higher energy via the Thomson scattering. The microwave spectrum is distorted. Photons, passing through the plasma at the temperature of about 10 keV (1.116×10^8 K), will be inverse Compton scattered. The percentages of photons that are scattered receive a boost that is proportional to $\frac{k_{\rm B}T_{\rm e}}{m_{\rm e}c^2}$ and the spectrum is distorted because photons transit from lower to higher frequencies. The thermal effect modifies the background black body spectrum of a very small amount. For $h\nu = kT$, that is at frequency of about 56 GHz for $T \sim 10$ keV, the effect is about:

$$\Delta T = 2T \int_{-\infty}^{+\infty} \sigma_{\rm T} n_{\rm e} \frac{k_{\rm B}T_{\rm e}}{m_{\rm e}c^2} dr \sim 0.5 \text{ mK},$$

for a bright cluster.

Details as a function of frequency in Fig. 9, where to evidence both the background spectrum and the distortion due to the presence of the cluster‘s plasma the two curves have a different normalization: the Black Body peak is a factor 10^5 more intense than the peak of the perturbation. The strength of the distortion is clearly related to the number of scatters and therefore to the optical depth of the plasma. In a more accurate way it is strongly dependent from the Compton parameter $y = \int \sigma_{\rm T} n_{\rm e} \frac{k_{\rm B}T_{\rm e}}{m_{\rm e}c^2} dr$ that is also related to the mass of the cluster. The distortion $\frac{\Delta T_{\rm SZ}}{T_{\rm MWB}} = f(\frac{h\nu}{kT})\, y$ is independent of the redshift! The survey by the South Pole Telescope (STP, Carlstrom, 2011) detected 677 cluster candidates over an area of 2500 square degrees (Bleem, 2015) – Fig. 10. Essentially the detection is based on a negative peak algorithm detection and the measured parameter (the negative peak or

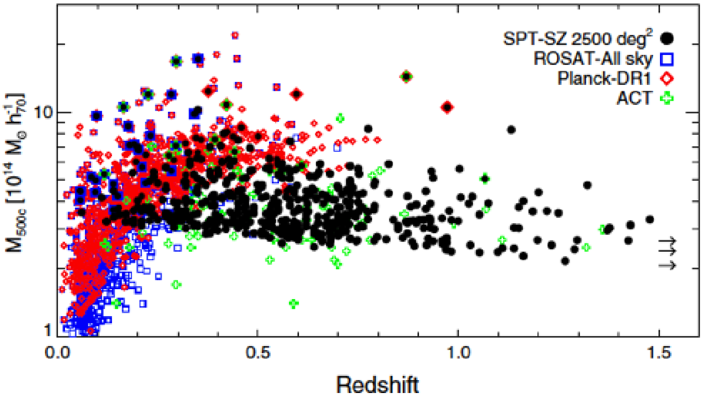

Fig. 10: From Bleem (2015). The SPT data provide a nearly mass limited sample.

better the integrated Comptonization parameter) is calibrated on the cluster mass using simulations, Vanderlinde (2010). The SZ surface brightness is independent of redshift (function only of $n_e T_e$) so that the SZ surveys are and excellent technique to discover clusters at any redshift. The SZ flux a robust proxy for the total cluster mass. All of this is fundamental for DM cosmology studies.

These surveys may lead us to map the DM as a function of redshift in the presence of gas. As for the accuracy of the SZ effect mass estimate ... “In conclusion we find statistically consistency between masses derived from WL data and those derived from SZ and X-ray data at the 20% level. Bradač et al. (2006) Clowe et al. (2004–2006) discovered that the plasma in the interacting clusters $1E067-558$ does not coincide with the distribution of galaxies and with the distribution of matter (DM) as measured by weak lensing. The interacting clusters differ in redshift of about $700\,\mathrm{km\,s^{-1}}$ and are separated on the perpendicular to the line of sight by 0.72 Mpc. The distribution of galaxies on both clusters agrees with the distribution of mass as estimated by the weak lensing observations. The collision where the smaller cluster crosses the main structure perturbs only the gas, galaxies and dark matter are transparent to each other, shifting the gas of the main cluster toward the smaller structure while the gas of the latter presents the classic shape of a bow shock. As discussed by Bradač et al. (2006) this is likely the best proof of the existence od Dark Matter (nowadays we have many observations of colliding clusters, see the presentation of lecture III). This leads us to the question on whether the distribution of baryons and DM differs also in not colliding clusters. In other words the distribution of cluster parameters as a function of the distance from the centre is a matter that needs to be known in detail.

Vikhlinin et al. (2006) use a Chandra sample of 13 low redshift relaxed galaxy clusters to study the total density and the gas profile. As shown in Fig. 11 for large radii the gas profile tends to get closer to the total density profile, that is the ratio of the baryonic mass to the DM mass changes. Vikhlinin et al. (2006) state: “We present accurate measurements of the gas mass fraction as a function of radius. We observe strong systematic variations of f_{gas} both with radius and with cluster mass. The gas fractions within r_{2500} are significantly lower than the universal

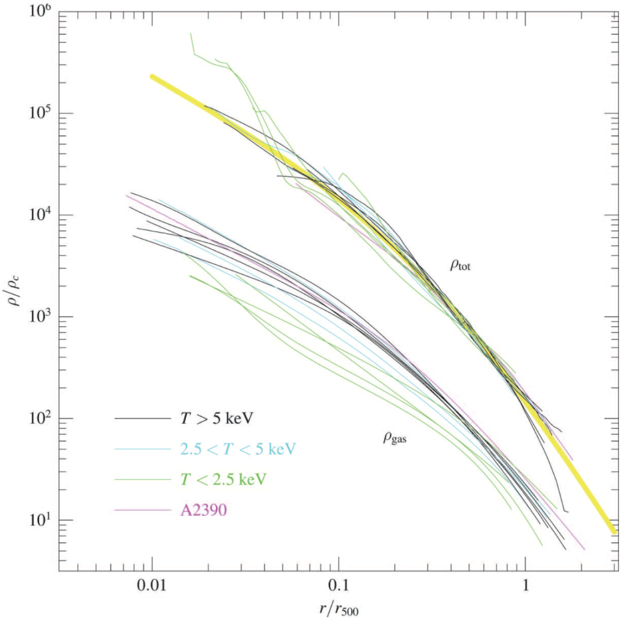

Fig. 11: Total density profiles and gas density profiles for 13 clusters. The yellow continuous line is the NFW profile. From Vikhlinin et al. (2006)

baryon fraction suggested by the CMB observations. However, the trends become weaker and the absolute values of fgas are closer to the universal value at $r > r_{2500}$". The gas fraction increases with radius as power law of the over-density. On average $\left\langle \frac{f_{2500}}{f_{500}} \right\rangle = 0.84$. Not only we need to estimate the dark matter mass possibly as a function of scale and redshift, but also the variation of the baryonic to DM mass. We would need to approach, to understand the distribution of matter – baryonic & non – what is generally known as the missing baryons problem.

The baryon fraction correlates with cluster mass, Vikhlinin et al. (2009a,b), according to the following relation where M_{15} is the cluster total mass, M_{500}, in units of $10^{15}h^{-1}M_\odot$: $f_g(h/0.72)^{15} = 0.125 + 0.037 \log M_{15}$. The ratio M_{500}/M_* varies from ~ 130 to ~ 280. The analysis of the baryonic mass has been analyzed in detail also by Chiu (2016) in the context of the SPT-SZ survey and for cluster-mass $\sim 6 \times 10^{14} M_\odot$ at a median redshift $z = 0.9$. In this case the masses of the galaxies have been estimated using 6 passbands and the models by Bruzual & Charlot (2003) while the X-ray observations have been based on 11 clusters on Chandra and on XMM for the other 3. They find:

$$f_* \equiv \frac{M_*}{M_{500}} = 1.1 \pm 0.1(\%),$$

$$f_{\rm ICM} \equiv \frac{M_{\rm ICM}}{M_{500}} = 9.6 \pm 0.5(\%),$$

$$f_{\rm collapsed} \equiv \frac{M_*}{M_{\rm b}} = 10.7 \pm 1.1(\%),$$

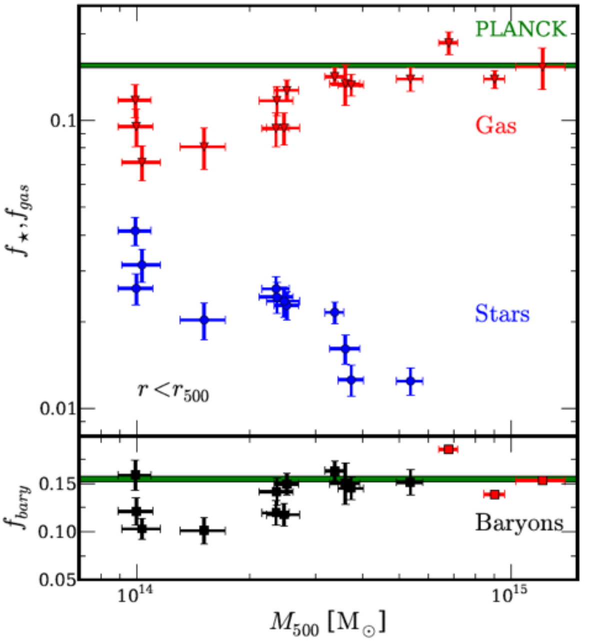

Fig. 12: Stellar, gas, and total baryon fraction as a function of cluster mass. The stellar data (blue circles) include the deprojection correction. The best fit to this stellar data is $f_{*,3D} \propto M_{500}^{-0.48\pm0.04}$, confirming evidence for a steep slope. The gas fractions (red triangles) are plotted for all systems including the massive clusters from Vikhlinin et al. (2006) for which we lack stellar fractions. The opposing trends for gas and stellar baryon fractions imply strongly decreasing star formation efficiency with increasing cluster mass. The total baryon fractions in the lower panel are plotted as black squares for all systems with stellar and gas data. As red squares the gas fractions for the most massive clusters that lack stellar data; these are formally lower limits on the total baryon fractions, with stellar baryon fractions expected to be at the level of $f_*\, 0.01 - 0.02$. The total baryon fractions are consistent with a weak, but statistically significant dependence of the total baryon fraction upon $M_{500}(f_{\rm bary} \propto M_{500}^{0.16\pm0.04})$ over the mass range where we have both stellar and gas measurements. This trend is driven primarily by the lower mass systems, for which we also see evidence of large scatter in the total baryon fraction. The weighted mean baryon fraction for the Gonzalez et al. (2013) Gonzalez sample at $M > 2\times10^{14} M_\odot$, $f_{\rm bary} = 0.136 \pm 0.005$ is 18% below the Universal value. From Gonzalez et al. (2013).

$$f_{\rm b} \equiv \frac{M_{\rm b}}{M_{500}} = 10.7 \pm 0.6(\%),$$

check

$$\Rightarrow f_{\rm c} \times f_{\rm b} = \frac{M_*}{M_{\rm b}} \frac{M_{\rm b}}{M_{500}} = f_* = 0.107 \times 0.107 = 0.011.$$

According to Gonzalez et al. (2013), Fig. 12, at large masses we are getting close to the cosmological values and the matter now is to understand the running of the various components of matter as a function of the total mass (baryonic and DM) and as a function of the distance from the centre of the cluster accounting for various heating and cooling mechanisms of the gas (Young et al., 2011; Schaller et al., 2015).

The challenge is the understanding of the interplay between DM and baryons and the discovery of where are the “missing” baryons. If we started everywhere from in the Universe from the same percentage of DM and baryons then we have to identify in which state the baryons are and which mechanism made them undetectable so far. The ratio baryons to Dark Matter is now on very robust grounds!

References

Binney, J., Tremaine, S., Galactic Dynamics, Princeton University Press, New Jersey (1987)

Bleem, L. E. e. a., *ApJS* **216**, 27 (2015)

Boldt, E., McDonald, F. B., Riegler, G., Serlemitsos, P., *Phys. Rev. Lett.* **17**, 447 (1966)

Bradač, M., et al., *ApJ* **652**, 937 (2006)

Briel, U. G., Henry, J. P., Boehringer, H., *A&A* **259**, 2, L31 (1992)

Bruzual, G., Charlot, S., *MNRAS* **344**, 4, 1000 (2003)

Carlberg, R. G., Yee, H. K. C., Ellingson, E., *ApJ* **478**, 2, 462 (1997)

Carlberg, R. G., et al., *AJ* **462**, 32 (1996)

Carlstrom, J. E. e. a., *PASP* **123**, 568 (2011)

Cavaliere, A., Fusco-Femiano, R., *A&A* **49**, 137 (1976)

Cavaliere, A., Fusco-Femiano, R., *A&A* **70**, 677 (1978)

Cavaliere, A., Fusco-Femiano, R., *A&A* **100**, 194 (1981)

Chandrasekhar, S., An introduction to the study of stellar structure, University of Chicago Press (1939)

Chincarini, G., *Nature* **272**, 515 (1978)

Chincarini, G., *Proceedings of the 13th Marcel Grossmann Meeting (Stockholm 1-7 July 2012)* (2013)

Chincarini, G., Rood, H. J., Welch, G. A., *MNRAS* **170**, 441 (1975)

Chincarini, G. L., Haynes, M. P., Haynes, M. P., *AJ* **84**, 1500 (1979)

Chiu, I. e. a., *MNRAS* **455**, 258 (2016)

Colless, M., Dunn, A. M., *AJ* **458**, 435 (1996)

Danese, L., de Zotti, G., di Tullio, G., *A&A* **82**, 322 (1980)

Dekel, A., Shaham, J., *A&A* **74**, 186 (1979)

Eddington, A. S., *MNRAS* **76**, 525 (1916)

Felten, J. E., Gould, R. J., Stein, W. A., Woolf, N. J., *ApJ* **146**, 955 (1966)

Geller, M. J., Diaferio, A., Kurtz, M. J., *ApJ* **517**, L23 (1999)

Giacconi, R., et al., *ApJ* **230**, 540 (1979)

Gonzalez, A. H., Sivanandam, S., Zabludoff, A. I., Zaritsky, D., *ApJ* **778**, 14 (2013)

Goobar, A. e. a., *Science* **356**, 291 (2017)

Governato, F., et al., *MNRAS* **307**, 949 (1999)

Hoekstra, H., *MNRAS* **379**, 317 (2007)

Hughes, J. P., *AJ* **337**, 21 (1989)

Humason, M. L., Mayall, N. U., Sandage, A. R., *AJ* **61**, 97 (1956)

Jaffe, W. J., Perola, G. C., *A&A* **26**, 423 (1973)

Jauzac, M. e. a., *MNRAS* **463**, 3876 (2016)

Kelly, P. L. e. a., *Science* **347**, 1123 (2015)

Kent, S. M., Gunn, J. E., *AJ* **87**, 945 (1982)

King, I., *AJ* **67**, 471 (1962)

King, I. R., *AJ* **174**, L123 (1972)

Kubo, J. M., et al., *ApJ* **671**, 1466 (2007)

Landau, L., Lifshitz, E., Mechanics III Edition, Elsevier Publisher (1976)

Limber, D. N., *ApJ* **131**, 168 (1960)

Liu, F. K., *MNRAS* **281**, 1197 (1996)

Łokas, E. L., Mamon, G. A., *MNRAS* **343**, 401 (2003)

Mayall, N. U., *Annales d'Astrophysique* **23**, 344 (1960)

Merritt, D., *AJ* **313**, 121 (1987)

Miley, G. K., Perola, G. C., van der Kruit, P. C., van der Laan, H., *Nature* **237**, 269 (1972)

Navarro, J. F., Frenk, C. S., White, S. D. M., *ApJ* **462**, 563 (1996)

Ogorodnikov, K. F., Dynamics of stellar systems, The Macmillan Company (1965)

Omer, J., Guy C., Page, T. L., Wilson, A. G., *AJ* **70**, 440 (1965)

Raychaudhury, S., *Nature* **342**, 251 (1989)

Rood, H. J., Page, T. L., Kintner, E. C., King, I. R., *AJ* **175**, 627 (1972)

Scaramella, R., et al., *Nature* **338**, 562 (1989)

Schaller, M., et al., *MNRAS* **452**, 343 (2015)

Sharon, K., et al., *ApJ* **746**, 161 (2012)

Sunyaev, R. A., Zeldovich, Y. B., *Comments on Astrophysics and Space Physics* **4**, 173 (1972)

Turnrose, B. E., Rood, H. J., *AJ* **159**, 773 (1970)

Vanderlinde, K. e. a., *ApJ* **722**, 1180 (2010)

Vikhlinin, A., et al., *ApJ* **640**, 691 (2006)

Vikhlinin, A., et al., *ApJ* **692**, 1033 (2009a)

Vikhlinin, A., et al., *ApJ* **692**, 1060 (2009b)

Young, O. E., Thomas, P. A., Short, C. J., Pearce, F., *MNRAS* **413**, 691 (2011)

Zwicky, F., *Helvetica Physica Acta* **6**, 110 (1933)

Zwicky, F., Morphological astronomy, Springer Verlag (1957)

Zwicky, F., Herzog, E., Wild, P., *Pasadena: California Institute of Technology* (1961-1968)