Improving NASA/IPAC Extragalactic Database Redshift Calculations

Anthony Carr[ID] and Tamara Davis[ID]

School of Mathematics and Physics, The University of Queensland, QLD 4072 Australia; anthony.carr@uq.net.au

Abstract

The NASA/IPAC Extragalactic Database (NED) is an impressive tool for finding near-exhaustive information on millions of astrophysical objects. Here, we outline a small systematic error that occurs in NED because a low-redshift approximation is used when making the correction from redshifts in the heliocentric frame to the cosmic microwave background (CMB) rest frame. It means that historically NED systematically misreported the values of CMB-frame redshifts by up to $\sim 10^{-3}z$ (about 0.001 at redshift of 1). This is a *systematic* error, and therefore the impact on applications requiring precise redshifts has the potential to be significant—for example, a systematic redshift error of $\sim 10^{-4}$ at *low* redshift could resolve the Hubble tension. We have consulted with the NED team and they are updating the software to remove this systematic error so these corrections are accurate at all redshifts. Here, we explain the changes and how they impact the redshift values NED currently reports.

Unified Astronomy Thesaurus concepts: Galaxies (573); Astronomy databases (83); Cosmology (343)

1. Introduction

Redshifts are absolutely fundamental to cosmology, and have been so *easy* to precisely measure for so long (e.g., Slipher 1921; Zwicky 1933; Tonry & Davis 1979) that the uncertainties in redshifts are often neglected, and certain approximations generally remain in use and some forms of potential systematic error remain unaccounted for (Wojtak et al. 2015; Davis et al. 2019). As other measurements become more precise, systematic errors in redshift are becoming more relevant. Perhaps the most topical and applicable problem we face in cosmology is the tension in the Hubble constant (H_0) between local and global measurements (Freedman et al. 2019; Riess et al. 2019; Planck Collaboration et al. 2020; Wong et al. 2020). Systematic errors in redshift as small as 10^{-4} could go partway in reconciling (or worsening) the H_0 controversy (Davis et al. 2019) since very small systematic redshift offsets—especially if they occur at low redshift—can have a large impact on the inferred cosmological parameters.

The NASA/IPAC Extragalactic Database (NED)[1] keeps track of many redshifts used for cosmology. In this paper we describe improvements to the computations needed to remove a systematic error in the reporting of redshifts corrected to the cosmic microwave background (CMB) reference frame. We have also worked with NED staff to update the help pages so those performing heliocentric corrections with NED or other data are aware of the correct method that is accurate at all redshifts. While the error is small and will have been negligible for most applications (e.g., Glanville et al. 2021), it is important to correct since systematic effects could propagate out to inferred cosmological parameters (Colless 1995; Wojtak et al. 2015; Calcino & Davis 2017; Chaves-Montero et al. 2018; Steinhardt et al. 2020).

We begin by clarifying the theory of redshifts and velocities, and discussing which dipole should be used for heliocentric-to-CMB corrections (Section 2). We then provide plots of NED redshifts before and after correction to demonstrate the size of the systematic errors (Section 3).

2. Theory

NED currently uses an additive approximation to combine different sources of redshift. In this approximation, $z_{\rm obs}$ is treated as an additive combination of the redshift due to cosmological expansion $z_{\rm CMB}$ and the redshift due to our Sun's (or more correctly the solar system's barycenter) peculiar motion $z_{\rm p,Sun}$ with respect to the CMB. When this approximation is used we give $z_{\rm CMB}$ the superscript "+", so

$$z_{\rm obs} = z_{\rm CMB}^{+} + z_{\rm p,Sun}. \tag{1}$$

However, the correct way to combine redshifts is to multiplicatively combine factors of $(1 + z)$. This is due to the definition of redshift $(1 + z)$ as the ratio of observed to emitted wavelength (see, e.g., Davis et al. 2019, Section 2.3 for an explanation). When this precise equation is used we give $z_{\rm CMB}$ the superscript "×", so

$$(1 + z_{\rm obs}) = (1 + z_{\rm CMB}^{\times})(1 + z_{\rm p,Sun}). \tag{2}$$

Expanding this second equation out and subtracting Equation (1),

$$z_{\rm CMB}^{+} = z_{\rm CMB}^{\times} + z_{\rm CMB}^{\times} z_{\rm p,Sun}. \tag{3}$$

Thus, the difference between using the additive approximation and the correct multiplicative equation is exactly $z_{\rm CMB}^{\times} z_{\rm p,Sun}$. A summary of symbols can be found in Table 1.

The redshift due to our Sun's motion of about 369.82 ± 0.11 km s^{-1} (Planck Collaboration et al. 2020) peaks in the direction of motion at $z_{\rm p,Sun} = 0.0012336 \pm 0.0000003$. Therefore at low redshift, when $z_{\rm CMB}^{\times}$ and the approximate $z_{\rm CMB}^{+}$ are small, their difference is negligible. By $z \sim 1$, however, the error is on the order of 10^{-3}, which is an order of magnitude larger than most reported statistical errors in spectroscopic redshifts. The systematic error remains even when using photometric redshifts, which generally have uncertainties $\gtrsim 10^{-3}$—an underlying systematic error by its definition will not be mitigated by averaging over many measurements.

Thankfully, despite this error growing larger with redshift, the impact on H_0 measurements from standard candles shrinks; these measurements are more strongly impacted by low-redshift

---

[1] https://ned.ipac.caltech.edu/

2

Table 1
Definitions of Symbols and Their Physical Description

Symbol	Name	Description
Generic recession velocity; redshift	v_r; z_{CMB}	Recession velocity (that appears in the Hubble–Lemaître law) and redshift due to the expansion of the universe. The recession velocity is related to the recession redshift through Equation (5).
Generic peculiar velocity; redshift	v_p; z_p	Peculiar velocity and redshift due to motions that deviate from the Hubble flow. Can refer both to our peculiar motion and the peculiar motion of objects we are observing.
Observed redshift	z_{obs}	The redshift of an object in the heliocentric frame, as given by NED. Contains contributions from the object's and our own peculiar motion relative to the CMB, and the object's recession. (Note this is not strictly observed redshift, as the velocity of the Earth relative to the Sun is generally removed.)
CMB-frame redshift derived multiplicatively	$z_{CMB}^{\times}$	The redshift of the object after correcting for our peculiar motion relative to the CMB using Equation (2). May still have contributions from the object's peculiar motion. This is the correct CMB-frame redshift.
CMB-frame redshift derived additively	z_{CMB}^{+}	The approximate redshift of the object after correcting for our peculiar motion relative to the CMB using Equation (1). This is a poor approximation at high z.
Sun's peculiar redshift/velocity	$v_{p,Sun}$; $z_{p,Sun}$	An object's peculiar velocity and peculiar redshift due to the peculiar velocity of the Sun in the direction of the object, relative to the CMB. Note that when $v_{p,Sun}$ is positive, $z_{p,Sun}$ is negative.
Sun's velocity toward dipole	$v_{p,Sun}^{max}$	The magnitude of the Sun's velocity in the CMB frame, related to $v_{p,Sun}$ through Equation (6).
Separation from dipole	α	The on-sky angular separation of an object from the direction of the CMB dipole.

Note. In addition to the base symbols, superscripts of "Planck" and "COBE" will often be added when referring to a parameter whose value changes depending on the dipole value used.

systematic errors (Wojtak et al. 2015; Davis et al. 2019). Nevertheless, it remains important to remove entirely as it will impact other measurements such as whether dark energy is time varying.

2.1. Velocities

Velocities are a complex topic in cosmology because they are coordinate dependent and we generally use coordinates that violate some of our naïve expectations of velocity behavior. This is in contrast to redshifts, that are observable, dimensionless ratios and thus unambiguous.

The low-redshift approximation $z = v/c$ is appropriate for both peculiar velocities and (small) recession velocities. However, the two types of velocities are quite different. Peculiar velocities, v_p, are measured with respect to a local Minkowski (special relativistic) frame, and thus the redshift induced by a velocity obeys the special relativistic Doppler shift formula,

$$1 + z_p = \sqrt{\frac{1 + v_p/c}{1 - v_p/c}}. \tag{4}$$

On the other hand, the recession velocity that appears in the Hubble–Lemaître law, v_r, cannot be described in any inertial frame. It can be derived from the Friedmann–Lemaître–Robertson–Walker metric as the purely expansion component of the time derivative of proper distance, D (Davis & Lineweaver 2004). The other component, from changes in comoving coordinate caused by gravitation, are the peculiar velocities above. Indeed for $v_r = H_0 D$ to be true at any D, v_r must be allowed to exceed the speed of light. (An analogous situation occurs with superluminal infall velocities of *space* inside a black hole, which is why light cannot escape.) This is a standard result, which is a requirement for homogeneity and isotropy, and does not violate relativity. Thus recession velocities do not obey the special relativistic Doppler shift formula, but are instead calculated by integrating over the expansion rate of the universe during the photon's propagation (Harrison 1993),

$$v_r = c \int_0^{z_{\mathrm{CMB}}} \frac{dz}{E(z)}, \tag{5}$$

where the dimensionless Hubble parameter $E(z) = H(z)/H_0 = \sqrt{\sum \Omega_i (1+z)^{3(1+w_i)}}$ (Peebles 1993; Ryden 2003) with $H(z)$ being the Hubble parameter at the time of emission from a galaxy at redshift z, Ω_i being the normalized density of component i of the universe, and w_i being that component's equation of state. It is interesting to note that the recession velocity as a function of *redshift* is actually independent of H_0.

The total velocity, v_t, of an object moving away from us can be accurately calculated as $v_t = v_r + v_p$ (Davis & Lineweaver 2004). The common additive approximation for redshifts comes from dividing this equation by c and applying the low-z approximation $z = v/c$. However, given the current precision of other measurements, this will be a poor approximation for any redshift higher than $z \sim 0.01$.

2.2. Converting Heliocentric Redshifts to the CMB Frame

To correct z_{obs} to the CMB frame requires finding $z_{p,\mathrm{Sun}}$. The motion of the solar system adds a peculiar redshift to every object. That peculiar redshift is highest for objects that are in or directly opposite the direction of motion, resulting in a maximum $z_{p,\mathrm{Sun}}$ of order 10^{-3}. The effective radial peculiar velocity of any object is the projection of the Sun's velocity vector on the position vector of the object,

$$v_{p,\mathrm{Sun}} = \boldsymbol{v}_{p,\mathrm{Sun}} \cdot \hat{\boldsymbol{n}}_{\mathrm{obj}} = v_{p,\mathrm{Sun}}^{\max} \cos \alpha, \tag{6}$$

where $v_{p,\mathrm{Sun}}$ represents the projection of the Sun's peculiar velocity along the line of sight to the object, $\hat{\boldsymbol{n}}_{\mathrm{obj}}$ is the object's position vector, and α is the angle separating the dipole direction and the object.[2]

We define both recession and peculiar velocities to be positive when they are moving away from us. Note that in the case of our own motion, a positive velocity gives a negative redshift (blueshift) because we are moving toward the object being observed.

Since the Sun's velocity is small (order of 10^2 km s^{-1}) compared to c, the low-z approximation $z_{p,\mathrm{Sun}} \approx -v_{p,\mathrm{Sun}}/c$ can be used. However, there is negligible computational disadvantage to using the full special relativistic calculation Equation (4),

$$z_{p,\mathrm{Sun}} = \sqrt{\frac{1 + (-v_{p,\mathrm{Sun}})/c}{1 - (-v_{p,\mathrm{Sun}})/c}} - 1. \tag{9}$$

The minus signs before $v_{p,\mathrm{Sun}}$ have been left explicit to ensure that a zero separation ($\alpha = 0$ in Equation (6)) results in the object appearing blueshifted due the our positive velocity toward it. For the purposes of the plots and equations in this paper, we emphasize that we use the approximation purely for clarity and so our comparisons with NED are cleaner (since it makes negligible difference relative to the size of the other discrepancies).

2.3. Multiple CMB Dipole Measurements

The correction requires the direction and magnitude of the CMB dipole, which was measured in 2018 by the Planck Collaboration (Planck Collaboration et al. 2020) to be in the direction of $(l, b) = (264.021° \pm 0.011°, 48.253° \pm 0.005°)$ with a velocity $v_{p,\mathrm{Sun}}^{\mathrm{Planck}} = 369.82 \pm 0.11$ km s^{-1}. However, NED currently uses the older measurement of the dipole by the Cosmic Background Explorer (COBE) satellite (Fixsen et al. 1996), $(l, b) = (264.14° \pm 0.30°, 48.26° \pm 0.30°)$ with a velocity $v_{p,\mathrm{Sun}}^{\mathrm{COBE}} = 371 \pm 1$ km s^{-1}.[3]

We compare the different combinations of dipole and heliocentric correction form below. Performing the multiplicative correction, Equation (2), with the modern dipole

---

[2] The formula for the angular separation, α, in terms of galactic longitude and latitude, (l, b), is

$$\cos \alpha = \sin(b)\sin(b_0) + \cos(b)\cos(b_0)\cos(l - l_0), \tag{7}$$

where (l_0, b_0) correspond to the dipole direction. An equivalent calculation that is more numerically stable at very small and large separations is the Vincenty formula (Vincenty 1975), for which we use *Astropy*'s SkyCoord implementation. Setting $\Delta l = l - l_0$, it reads as

$$\tan \alpha = \frac{\sqrt{(\cos b \sin \Delta l)^2 + (\cos b \sin b_0 - \sin b \cos b_0 \cos \Delta l)^2}}{\sin b \sin b_0 + \cos b \cos b_0 \cos \Delta l}. \tag{8}$$

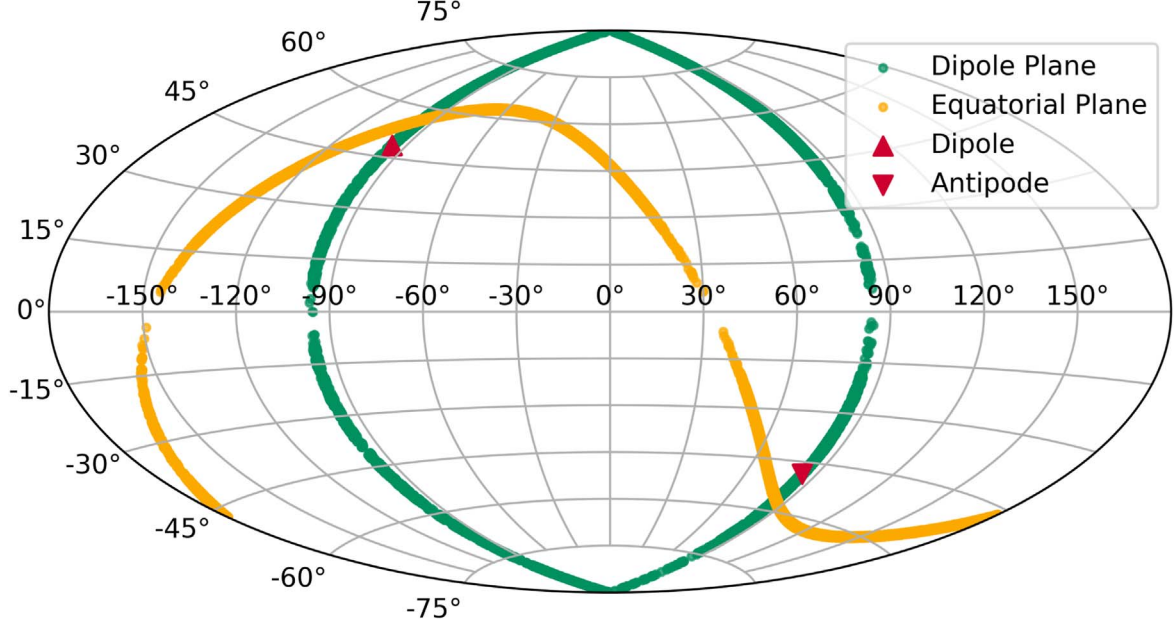[3] https://ned.ipac.caltech.edu/Documents/Guides/Calculators.

Figure 1. Data sample studied here, showing the equatorial subsample (yellow) and dipole plane (green) in galactic coordinates with the CMB dipole direction and its antipode marked with red triangles.

measurement is naturally the most accurate. However, we show explicitly NED uses the additive expansion, Equation (1), and the COBE dipole. In doing so, we demonstrate the size of the resulting systematic errors currently in NED, as well as in the cases of only updating the heliocentric correction form or dipole separately.

3. Testing NED Redshifts

To test NED redshifts we examine two different data selections. The first takes objects on a great circle encompassing the dipole direction and its antipode ($1.5°$ wide), and the second samples the equatorial plane ($0.6°$ wide), see Figure 1. As we detail below, adding an equatorial sample is helpful for understanding how errors in the heliocentric correction behave in different parts of the sky.

The combined data set was retrieved using NED in five parts, by splitting the dipole sample into two parts and the equatorial plane into three, for a total of nearly 197,000 objects. The equatorial sample had to be split into more parts to avoid the queries timing out, as there were nearly an order of magnitude more objects than in the dipole sample. At the time of retrieval, two XML VOTable files each were required for each data part; one contained the measured parameters, including z_{obs}, and the other contained derived data such as z_{CMB}, to compare to our calculations. The search criteria in addition to sky position included that the object be a NED classified extragalactic object —excluding *Galaxy groups* and *QSO groups*—with an available redshift. Despite requesting only extragalactic objects, some objects with negative redshift contaminated the sample. However, this does not pose a problem since all equations are equally valid for negative redshifts (even if the objects are too local to apply the heliocentric correction).

At the time of writing, the object search process has been altered so that only heliocentric velocities and redshifts can be queried in bulk.

3.1. Analytic Calculations Compared to Data

In Figures 2–5, we compare the CMB redshifts from NED to the CMB redshifts we calculate. We begin by confirming that NED currently uses the additive redshift calculation and the COBE dipole. In Figure 2, we show there is negligible difference in $z_{\mathrm{CMB}}^{+,\mathrm{COBE}} - z_{\mathrm{CMB}}^{\mathrm{NED}}$ since the absolute error only reaches 5×10^{-7}. This error comes purely from comparing the NED redshifts, that are rounded to six decimal places, with our non-rounded calculations.

We then improve on the NED redshifts by using the full multiplicative heliocentric correction and the Planck dipole. Using the multiplicative redshift calculations results in a redshift difference between our calculations and NED's on the order of 10^{-3} (Figure 3), while changing the dipole results in a redshift difference on the order of 10^{-6}–10^{-5} (Figures 4 and 5). Correcting the current NED CMB redshift calculation is therefore the most important change, while updating the COBE dipole is a valuable but subtle change.

Figure 3 shows the full difference between the Planck dipole with multiplicative redshift calculation, and the COBE dipole with additive redshift calculation. That is the difference between Equation (2) and Equation (1) that we calculate analytically to be

$$z_{\mathrm{CMB}}^{\times,\mathrm{Planck}} - z_{\mathrm{CMB}}^{+,\mathrm{COBE}} = -z_{\mathrm{p,Sun}}^{\mathrm{Planck}}(1 + z_{\mathrm{CMB}}^{\times,\mathrm{Planck}}) + z_{\mathrm{p,Sun}}^{\mathrm{COBE}}.$$
(10)

Recall that the dipoles differ in both magnitude and direction so there are the two separate expressions for the Sun's peculiar redshift $cz_{\mathrm{p,Sun}}^{\mathrm{Planck}} \approx -v_{\mathrm{p,Sun}}^{\mathrm{max,Planck}} \cos \alpha^{\mathrm{Planck}}$, and similar for COBE. We show in Figure 3 these analytic calculations (Equation (10)) by varying either z_{CMB} or α, as dashed lines. The excellent agreement when replacing $z_{\mathrm{CMB}}^{\mathrm{NED}}$ with $z_{\mathrm{CMB}}^{+,\mathrm{COBE}}$ confirms that indeed $z_{\mathrm{CMB}}^{\mathrm{NED}} = z_{\mathrm{CMB}}^{+,\mathrm{COBE}}$.
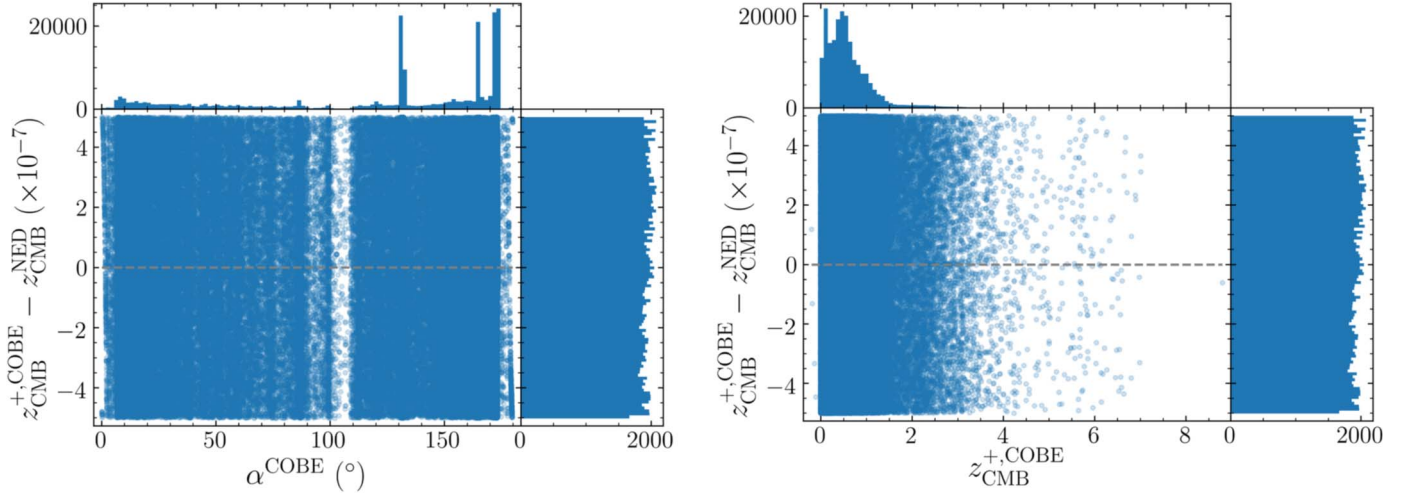
Figure 2. Difference between z_{CMB} calculated using the additive equation with the COBE dipole and the CMB redshift reported in NED. The agreement supports that $z_{CMB}^{NED} = z_{CMB}^{+,COBE}$ since the error arises purely from rounding. Left: the only trend is the clustering at certain separations caused by dense survey regions. Right: along the horizontal axis, the expected distribution of observed galaxies with redshift is evident.
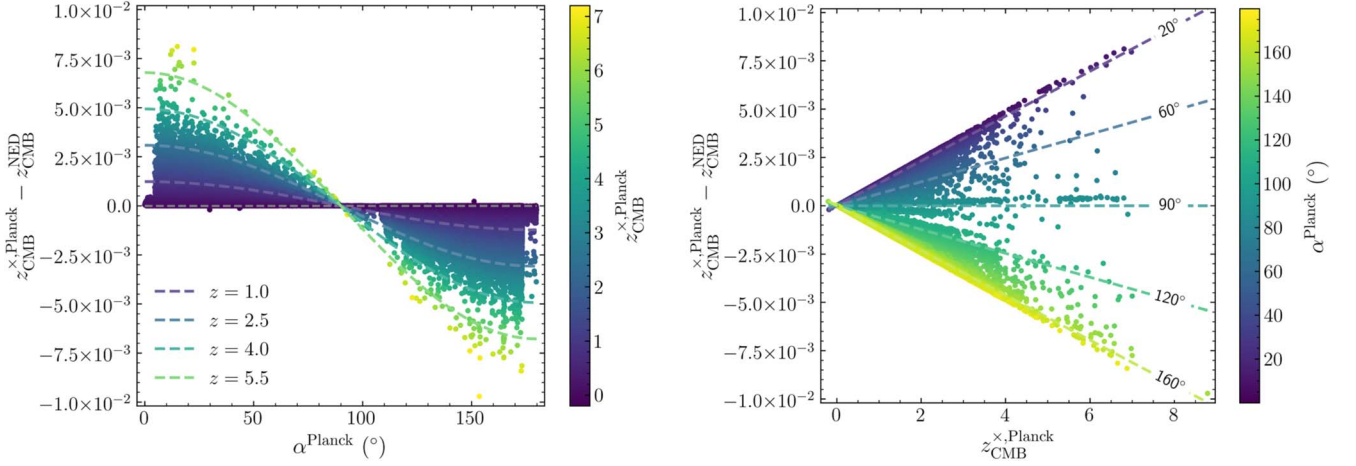

Figure 3. Difference between z_{CMB} calculated using the correct multiplicative equation for the Planck dipole and the CMB redshift reported in NED. The lines show the analytic prediction (Equation (10)) for the difference between the full calculation and NED values. Left: as a function of the objects' separation from the dipole direction, with the color bar representing the CMB-frame redshift. Right: as a function of the CMB redshift, with colors representing the angular separation from the dipole direction. (The contour colors have been lightened so as to be visible over the data.)
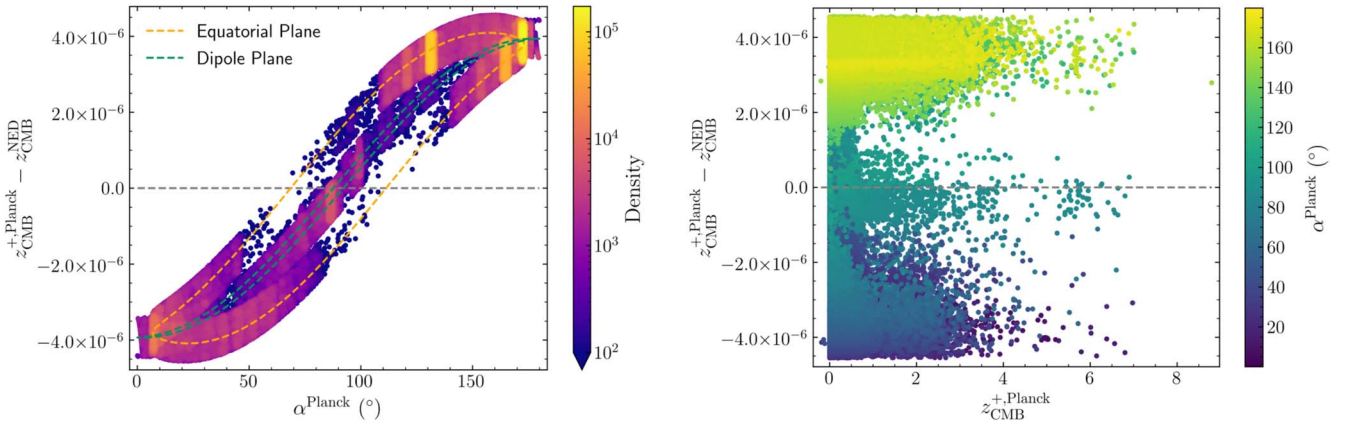

Figure 4. Difference in z_{CMB} using additive+Planck and NED. The expected error contours (dashed lines) now depend on exact sky position (the equatorial and dipole samples are now separated) as opposed to just 1D separation. Left: as a function of separation from the Planck dipole, with color now representing density of points to contrast the two data samples. Right: as a function of calculated CMB redshift, with color representing separation from the Planck dipole. The clear trend with separation angle (color bar) has been washed out slightly by the equatorial sample where maximum absolute no longer occurs at extremal (Planck) separations.
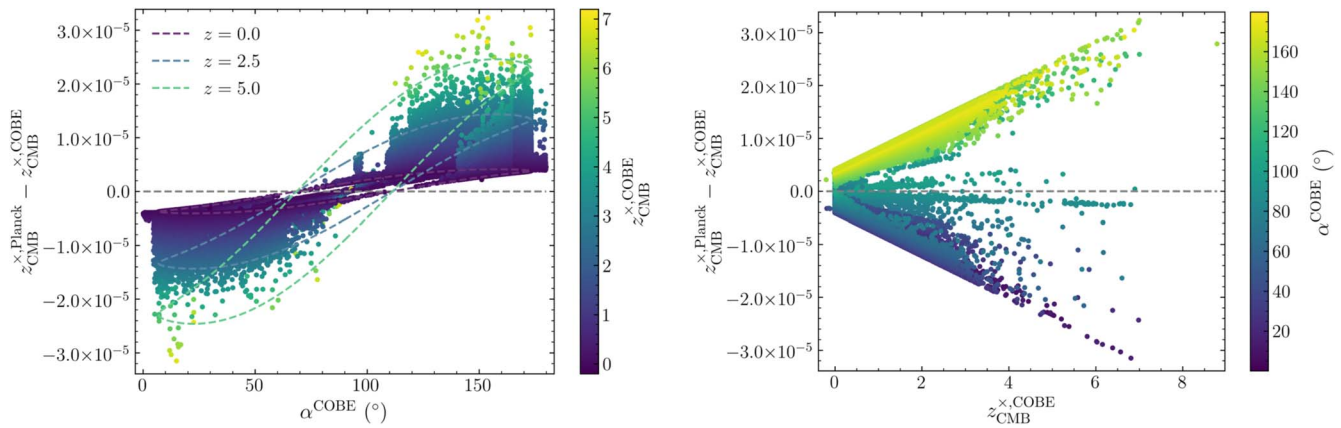
Figure 5. Comparison of full multiplicative heliocentric correction with the two dipoles. As discussed in the text, this essentially only differs from Figure 4 by a factor of z_{CMB}. Left: as a function of separation. There is excellent agreement with the theoretical contours (dashed lines). Note that the contours here are only for the equatorial sample. Right: as a function of z_{CMB} using the COBE dipole.

Figure 4 highlights only the difference caused by using the Planck dipole over COBE. For the purposes of this demonstration we calculate the CMB redshift using the additive approximation, but with the Planck dipole. The difference should therefore be

$$z_{CMB}^{+,Planck} - z_{CMB}^{+,COBE} = -z_{p,Sun}^{Planck} + z_{p,Sun}^{COBE}. \quad (11)$$

The more complicated shape shown in the left panel of Figure 4 is caused by this difference in two cosines (the different dipole directions) and very clearly shows the difference between the two data samples. The dipole plane sample has a small variation about the central sinusoid since the Planck dipole is very close to the COBE dipole, but the slight offset of roughly $0.1°$ causes a single value of α^{Planck} to correspond to two slightly different α^{COBE}. The equatorial sample has the larger variation because closest sky position to the dipoles is $6°$ away.

Figure 5 again compares the two dipoles, but this time using the correct multiplicative redshift calculation for both,

$$z_{CMB}^{\times,Planck} - z_{CMB}^{\times,COBE} = -z_{p,Sun}^{Planck}(1 + z_{CMB}^{\times,Planck})$$
$$+ z_{p,Sun}^{COBE}(1 + z_{CMB}^{\times,COBE}). \quad (12)$$

Comparing $z_{CMB}^{\times,Planck}$ to z_{CMB}^{NED} would be only imperceptibly different from Figure 3, so instead we opt to show the difference purely caused by the dipoles. This is essentially Figure 4 and Equation (11) with a multiplicative $1 + z_{CMB}$ factor. Therefore, using the full heliocentric correction with the outdated dipole measurement actually also multiplies the error. Note, however, that the sign of error when comparing dipoles is opposite to the additive versus multiplicative correction. The increased error here then cancels slightly more with the heliocentric correction form error (albeit still two orders of magnitude smaller).

4. Conclusion

The additive approximation to calculate the cosmological redshift was reasonable in the past when the correction was first implemented in NED, but this is no longer the case with the high redshifts of modern data. At a redshift of only 0.1, the error in correcting heliocentric redshifts to the CMB reference frame using the additive approximation is on the order of 10^{-4}, and this increases linearly with redshift (Figure 3).

A systematic error at that level would noticeably affect standard candle derivations of H_0 if it were at low redshift. The H_0 measurements are not likely to be strongly impacted, however, both because this correction is small at low z and because standard candle measurements come from many directions on the sky, so the positive and negative errors tend to cancel out. However, surveys that cover a small area of the sky are more likely to be biased by this effect. High-z supernova surveys, that tend to monitor small patches of sky, will be more significantly affected (both because of their high redshifts and the narrow α range). Correcting this is therefore most important for measuring features such as the equation of state of dark energy.

Finally, we acknowledge that the additive equations and COBE dipole are still used by many in the broader community beyond NED. We recommend dispensing with these approximations and using the full multiplicative equation to calculate CMB-frame redshifts, while also using the up-to-date Planck dipole. These are trivial changes to implement but has the potential to avoid significant errors in resulting analyses.

ORCID iDs

Anthony Carr ⓘ https://orcid.org/0000-0003-4074-5659
Tamara Davis ⓘ https://orcid.org/0000-0002-4213-8783

References

Calcino, J., & Davis, T. 2017, JCAP, 2017, 038
Chaves-Montero, J., Angulo, R. E., & Hernández-Monteagudo, C. 2018, MNRAS, 477, 3892
Colless, M. 1995, AJ, 109, 1937
Davis, T. M., Hinton, S. R., Howlett, C., & Calcino, J. 2019, MNRAS, 490, 2948

Davis, T. M., & Lineweaver, C. H. 2004, PASA, 21, 97
Fixsen, D. J., Cheng, E. S., Gales, J. M., et al. 1996, ApJ, 473, 576
Freedman, W. L., Madore, B. F., Hatt, D., et al. 2019, ApJ, 882, 34
Glanville, A., Howlett, C., & Davis, T. M. 2021, MNRAS, 503, 3510
Harrison, E. 1993, ApJ, 403, 28
Peebles, P. J. E. 1993, Principles of Physical Cosmology (Princeton Series in
    Physics) (Princeton, NJ: Princeton Univ. Press)
Planck Collaboration, Aghanim, N., Akrami, Y., et al. 2020, A&A, 641,
    A1
Riess, A. G., Casertano, S., Yuan, W., Macri, L. M., & Scolnic, D. 2019, ApJ,
    876, 85

Ryden, B. S. 2003, Introduction to Cosmology (Reading, MA (also London):
    Addison-Wesley)
Slipher, V. M. 1921, PA, 29, 128
Steinhardt, C. L., Sneppen, A., & Sen, B. 2020, ApJ, 902, 14
The Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018,
    AJ, 156, 123
Tonry, J., & Davis, M. 1979, AJ, 84, 1511
Vincenty, T. 1975, Surv. Rev., 23, 88
Wojtak, R., Davis, T. M., & Wiis, J. 2015, JCAP, 2015, 025
Wong, K. C., Suyu, S. H., Chen, G. C. F., et al. 2020, MNRAS, 498, 1420
Zwicky, F. 1933, AcHPh, 6, 110