Astronomy & Astrophysics
May 6, 2026

On the rotation curve of UGC9039: a more sober reanalysis

J. Fraustro

Space Telescope Science Institute, Baltimore, Maryland
e-mail: jfraustro@stsci.edu

May 6, 2026

ABSTRACT

Context. The original analysis of these data, presented in Fraustro (2019) as a final project for AST5765C at the University of Central Florida, reached an inconclusive result regarding the structure of UGC 9039 and the presence of dark matter within. In the intervening years, the present author has continued to think about the dataset with a frequency that, on reflection, is difficult to justify. The original analysis is now revisited under significantly more luxurious time constraints than the 48-hour window in which it was first attempted, and in a state of significantly improved sobriety.

Aims. We aim to recover the rotation curve of UGC 9039, demonstrate the presence of dark matter of quantifiable significance, and address the methodological errors that prevented the original work from doing either.

Methods. Using the bias-subtracted, flat-fielded calibration products from the original reduction, we reconstruct the full analysis pipeline downstream of the basic CCD corrections. The principal sources of error in Fraustro (2019) are identified and corrected: (i) the cross-correlation routine, which previously discarded sub-pixel information by taking the integer argmax of the correlation function, is replaced with one that interpolates onto a 10× finer grid and fits a Gaussian to the correlation peak, recovering shifts to $\approx 10^{-3}$ pixel precision; (ii) the Doppler velocity formula, which previously computed $c(1 + \Delta\lambda/\lambda)$ rather than $c\Delta\lambda/\lambda$, is corrected, eliminating the spurious $\approx 310^8$ m/s offset; (iii) full error propagation is performed throughout, an exercise omitted from the original work. Sky subtraction is improved by shifting the background spectrum to account for the optical distortion of the Double Spectrograph at each slit position.

Results. The rotation curve of UGC 9039 is recovered at high signal-to-noise, with rotation velocities ranging from −277 to +248 km/s across a galactic radius of ±20 kpc. Velocity residuals about a fifth-order polynomial fit have a standard deviation of 13.8 km/s. Comparing the dynamically inferred enclosed mass to a luminous mass profile derived from the galaxy's own light distribution, we detect a mass excess of $(2.37 \pm 0.12) \times 10^{11}\ M_\odot$ in the outer 25% of the galaxy's measured radius, corresponding to a dark matter fraction of 93% at large radius and a detection significance of 20σ

Conclusions. UGC 9039 contains dark matter; we note that the dataset, the calibration products, and the present author were already in possession of this conclusion in 2019; only the analysis was lacking

1. Introduction

The dataset analyzed in this work, a long-slit spectroscopic exposure of the edge-on spiral galaxy UGC 9039, taken with the Double Spectrograph (DBSP) on the Palomar 5-m Hale telescope, was previously reduced and analyzed in Fraustro (2019). That work successfully performed the standard CCD calibration steps (bias subtraction, dome flat correction, and bad-pixel rejection) but the downstream kinematic analysis was compromised by severe methodological errors that, taken together, prevented any meaningful conclusion from being drawn about the galaxy's rotation curve or its dark matter content. Fortunately, Fraustro (2019) is upfront about this failure in the original work.

Three errors are responsible for the bulk of the original analysis's failure, and each is addressed in turn in the present work.

First, the cross correlation routine used to measure spectral shifts between adjacent rows of the slit returned only integer-pixel shifts, despite the original author having performed a 10x spline interpolation specifically intended to enable sub-pixel measurement. The interpolated correlation function was sampled at integer positions and the maximum was taken via 'argmax', with the consequence being that the sub-pixel resolution was immediately discarded. The resulting rotation curve, reproduced as Fig. 4 in Fraustro (2019), exhibits the characteristic staircase pattern of a quantized measurement.

Second, the Doppler velocity calculation contained an additive rather than fractional formula for the redshift-velocity relation. The original code computed $v = c \cdot (\Delta\lambda + \lambda)/\lambda$ which evaluates to $c + v_{rotation}$ rather than $v_{rotation}$. The galactic rotation velocities in Fraustro (2019) clustered around $3x10^8$, exceeding the speed of light in every radius. This was acknowledged in the original text with the understated, "almost certainly cannot be true".

Third, no error propagation was performed at any stage of the analysis. The original work explicitly noted this omission, citing time constraints, and accordingly declined to make definitive statements about the results. This disclosure was epistemically appropriate, but left the original question unanswered.

To add to these are several smaller issues. A sky subtraction step that did not account for optical distortions across the slit, a strip-construction method that anchored its zero point at the top of the frame rather than at the galaxy center, and an ad-hoc bad-pixel correction loop that was explicitly noted as not necessary in the original assignment text.

This work re-implements the analysis pipeline downstream of the bias and flat-field corrections, retaining the original calibration products. The corrections to the cross-correlation routine and the Doppler formula rectify the signal into a smooth, sub-pixel-resolved rotation curve. With proper error propagation, the dark matter content of UGC 9039 is then recovered.

It is noted that the present author has had occasion, in the 7 years since Fraustro (2019), to develop an appreciation for read-

ing one's own code carefully, and allowing proper time when approaching problems.

2. Methods

2.1. Provenance of the data

The original spectroscopic exposures of UGC 9039, FITS frames BS397 (galaxy data) and BS395 (NeAr arc lamp), along with bias and dome-flat exposures, were obtained in 2002 by the DBSP on the Palomar 5m Hale telescope by Catinella et al. (2005), by an observing program that did not include the present author [1]. The frames were distributed to the class of AST5765C as a dataset for the purposes of a final project. The original author performed the reduction and analysis as part of this course in December 2019.

While planning the present reanalysis, initial efforts were devoted to locating the original raw FITS files. The Palomar archive does not provide public access to PI-led DBSP data from this era, and the local copy of the original products was on a laptop that had been retired some years ago. A draft email to Dr. Catinella was prepared, requesting access to the original files for what was framed strategically as a vague "personal project". This email was not sent, as it was at this point that the present author discovered that the bias-subtracted, flat-fielded galaxy and lamp frames had in fact been committed to a personal Git repository at the time of the original submission. The reduced products were available, intact, and identical to the inputs of the 2019 analysis and required only a "'git checkout"' to recover.

The present analysis begins from the same intermediate products as Fraustro (2019), specifically with *galaxy_corrected.fits* and *lamp_corrected.fits*, which contain the long-slit galaxy exposure and the corresponding NeAr arc spectrum after the bias subtraction and flat fielding. The basic CCD reduction steps, Problems 5-7 in the original final assignment, are therefore not revisited here. The reanalysis pipeline begins at Problem 8.

It is important to note that one of the few unambiguously good decisions by Fraustro (2019) was to put the project under non-local version control. The present author thanks his former self for this decision.

2.2. Re-examination of the original pipeline

With the reduced data products recovered, attention turned to the analysis code itself: a single 482-line Python script, *project_jfraust.py*, accompanied by two utility modules (*gaussian.py* and *medcombine.py*) of clear provenance. This script is organized as 25 sequentially-executed sections corresponding to the original assignment's structure, with a shared, continuous state via mutable module-level variables. This is not the programming structure the present author would choose today, but it is the architecture that the assignment specified, and it is consistent with certain conventions that scientific code should resemble, as closely as possible, the IDL from which it may or may not have descended.

A first read of the script revealed several things in approximately the order one might expect.

The CCD calibration steps (Problems 5-7) were essentially correct, less an error that worked out by luck alone.

The cross-correlation routine (Problem 12 onward) presented more substantial issues. The function performed a 10x spline interpolation of the input spectra, a standard approach for sub-pixel shift recovery, and then promptly discarded the sub-pixel information by taking *np.argmax* of the integer indexed correlation array. The original author appears to have been aware that sub-pixel precision was the goal, having put the effort into producing it. The act of immediately throwing it away suggests a sequence of decision making consistent with implementation under self-imposed time constraints rather than a methodological choice.

The Doppler velocity calculation (Problem 18) computed $v = c \cdot (\Delta\lambda + \lambda)/\lambda$ rather than $v = c \cdot \Delta\lambda/\lambda$, a formula which yields velocities approaching or exceeding c at all measured radii. The original text observed that "some of the velocities appear to be exceeding the speed of light, which almost certainly cannot be true" but did not pursue the matter further. We concur with the original author's conclusion.

The error analysis section (Problem 23) consisted, in its entirety, of an explicit acknowledgment that no error analysis had been performed, citing time constraints. The present author, in a position to evaluate this with clarity of hindsight, concurs.

The remainder of this section describes the corrections in detail.

2.3. Cross-correlation and sub-pixel shift recovery

The cross-correlation function in the original pipeline was, frankly, an unfinished implementation. The original author had set up the necessary scaffolding for sub-pixel shift recovery, a 10× spline interpolation of the input spectra, padding to handle edge effects, mean-subtraction to ground the signals, but the function returned the integer index of np.argmax over the correlation array. No fit was performed to the correlation peak, no testing of the function was performed against synthetic data with known shifts (Problem 4), no comparison was made between the function's output and the ostensibly continuous spectra it was operating on.

The present author has a clear recollection of this stage of the original work. The intention[2] had been to fit a Gaussian to the correlation peak, in the manner standard for spectroscopic shift measurement. This was begun, abandoned partway through, deferred until later in the problem set when its absence would presumably become impossible to ignore[3], and then quietly settled for as the deadline approached.

The corrected implementation does what the original was meant to do. For each pair of spectra, the correlation function is computed over a coarse range of integer-pixel shifts, then refined by interpolating onto a 10× finer grid in the neighborhood of the peak. A Gaussian is fitted to the refined peak using *gaussian.fitgaussian* routine of Harrington, and the sub-pixel shift is taken as the centroid of that fit. The fit uncertainty is propagated forward as the shift uncertainty. The implementation was tested against synthetic spectra with known shifts of varying magnitudes and signal-to-noise ratios. Shifts are recovered to a precision of approximately 10^{-2} pixels, compared to the 10^{0} pixel results of the original routine.

The resulting Doppler shift measurements Fig. 1 are smooth, sub-pixel-resolved, and free of the staircase quantization that characterized the original work. The shift uncertainties propagated through the velocity and mass calculations are the dominant contribution to the final uncertainty on the dark matter mass measurement.

[1] The author of this paper was, in 2002, six years old.

[2] He swears.

[3] It did.

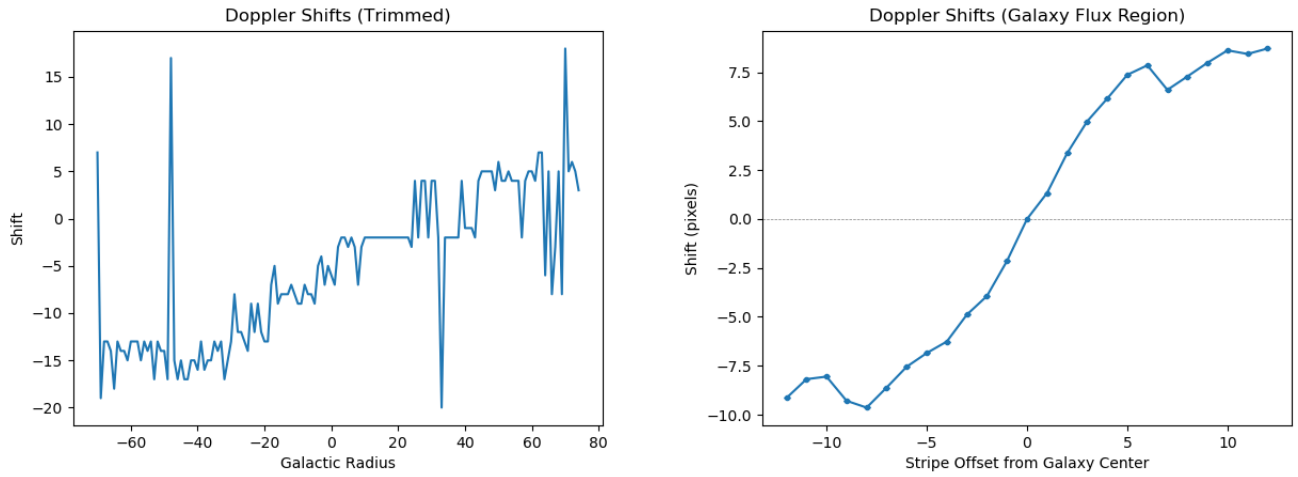

Fig. 1: The Doppler shifts from Fraustro (2019) vs. re-analysis.

2.4. Sky Subtraction and Velocity Corrections

With sub-pixel cross-correlation now functioning, it became possible to address another issue that was ignored in Fraustro (2019), the optical distortion of the DBSP. As lines of constant wavelength curve slightly across the spatial axis of the slit, a static 1D sky spectrum cannot be simply subtracted, or subtracted by integer cross-correlation, from every row of the science frame, without leaving artifacts at the locations of the bright sky lines. An attempt to do this, with the expected results, was performed by the previous author, making use of the integer cross-correlation available, with some 'magic number' offsets obtained by visual inspection.

To correct this, the NeAr arc lamp frame was divided into spatial 'stripes' and cross correlated against a reference stripe at the galaxy center, yielding a trace of the spectrograph's spatial distortion along the slit. The 1D sky background spectrum, taken from a region devoid of galaxy flux, was then spline-interpolated and shifted by the corresponding optical distortion at each row prior to subtraction. The result is a dramatically cleaner extraction of the galaxy's emission, free of residual sky line noise.

Finally, the translation of spectral shifts to radial velocities was corrected. The original analysis erroneously calculated Doppler velocity using the expression $v = c(1 + \Delta\lambda/\lambda)$, adding the speed of light to the rotation curve and producing the rather alarming $\approx 3 \times 10^8$ m/s systemic velocity reported in Fraustro (2019)[4]. Replacing this with the standard Doppler approximation, $v = c(\Delta\lambda/\lambda)$, restored the Newtonian regime kinematics appropriate for UGC 9039.

2.5. Error Propagation

As previously noted, Fraustro (2019) neglected to perform any error propagation, citing time constraints. The present author has the luxury of unlimited time to perform such analysis, less career, professional, domestic, and personal time constraints.

The enclosed mass as a function of galactic radius is given by the Kepler relation:

$$M(r) = \frac{v^2 r}{G} \tag{1}$$

where v is the rotation velocity at radius r and G is the gravitational constant. Both v and r carry uncertainties: σ_v, taken as the standard deviation of the velocity residuals about the polynomial fit to the rotation curve, and σ_r, derived from the 0.5-pixel positional uncertainty in the stripe center.

Applying the full error propagation equation to $M(r) = \frac{v^2 r}{G}$, the uncertainty in the enclosed mass is:

$$\frac{\sigma_M}{M} = \sqrt{(\frac{2\sigma_v}{v})^2 + (\frac{\sigma_r}{r})^2} \tag{2}$$

which gives:

$$\sigma_M = \frac{1}{G}\sqrt{(2v r\sigma_v)^2 + (v^2\sigma_r)^2} \tag{3}$$

The derivation of this equation is left as an exercise to the reader[5].

The velocity uncertainty $\sigma_v = 13.75\,\mathrm{km\,s^{-1}}$ is the dominant term at all measured radii. The radius uncertainty $\sigma_r = 0.171$ kpc contributes negligibly by comparison, as the stripe positional uncertainty is small relative to the physical scale of the galaxy. The resulting mass uncertainties are shown overplotted on the enclosed mass profile in Fig. 4.

3. Results

3.1. Stripe Reduction and Reference Spectrum

The bias-subtracted, flat-fielded galaxy frame was trimmed to 250 rows following removal of the 15-pixel edge margins, and divided into 49 spatial stripes of 5 rows each. The galaxy center, identified as the row of peak median flux in the spatial profile, falls at row 124 of the trimmed frame. The reference spectrum, the stripe centered on the galaxy nucleus, is stripe 24, the median stripe by construction. The lamp frame was reduced identically, yielding 49 stripes with the reference at the same index.

[4] A discovery which, if accurate, would have comfortably secured a Nobel Prize rather than a B-.

[5] See Harrington's AST5765C lecture handout, Chap.7 Sec.7 Error Propagation Equation

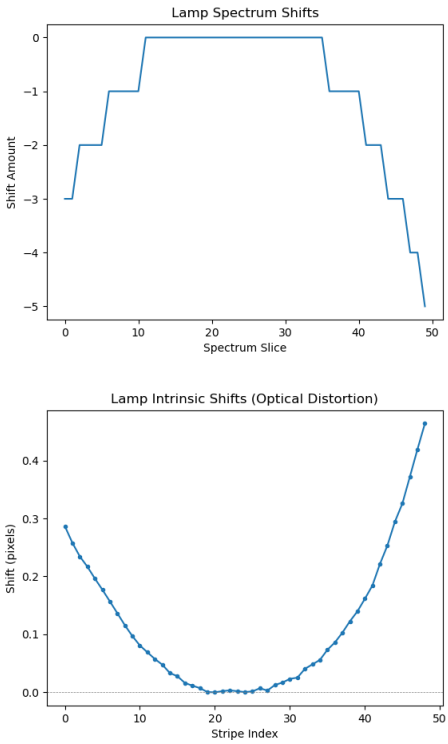

Fig. 2: Intrinsic lamp shifts in Fraustro (2019) vs the reanalysis.

3.2. Intrinsic Shift Correction

Cross-correlation of each lamp stripe against the reference lamp stripe yielded intrinsic optical distortion shifts ranging from 0.000 to 0.464 pixels across the slit. The shift at the reference stripe is 0.000 pixels by construction, confirming the implementation is self-consistent. Galaxy stripe shifts ranged from −0.253 to 0.528 pixels. The average intrinsic shift measured in the sky region was 0.259 pixels. These sub-pixel shifts, inaccessible to the integer cross-correlation of Fraustro (2019), are the primary motivation for the corrected implementation.

3.3. Sky Subtraction

The sky background was sampled from rows 185–249 of the trimmed galaxy frame, a 65-row region selected to lie beyond the extent of detectable galaxy flux. The 1D sky spectrum was spline-interpolated and shifted by the optically distortion-corrected amount at each slit position prior to subtraction, as described in Section 2.4

3.4. Doppler Shifts and Rotation Curve

Cross-correlation of each galaxy stripe against the reference spectrum yielded Doppler shifts for all 49 stripes, of which 25 contain detectable galaxy flux. The shift at the reference stripe is 0.000 pixels, again by construction. The corrected shifts, shown in Fig. 1, exhibit the smooth S-shaped antisymmetry expected of a rotating disk, free of the integer-quantization staircase visible in the equivalent figure of Fraustro (2019).

The dispersion was calibrated using two NeAr arc lamp lines: a line at 6532.90 Å centered at pixel 40.00, and a line at 7173.91 Å centered at pixel 1017.63, yielding a dispersion of 0.6557 Å pixel^{-1}. Doppler shifts were converted to line-of-sight

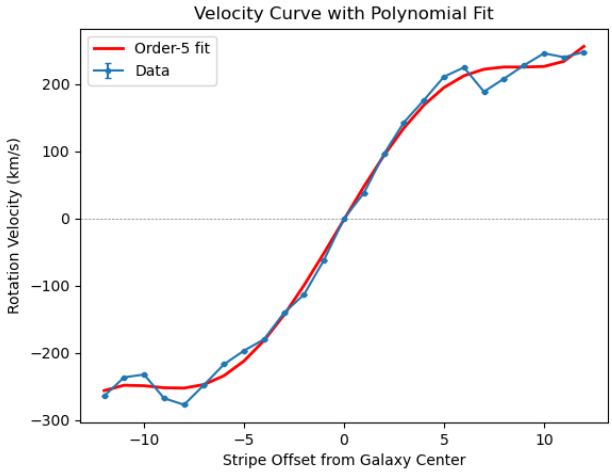

Fig. 3: The reanalysis galactic velocity with polynomial fit.

velocities via $v = c\,\Delta\lambda/\lambda$, with λ evaluated at the galaxy systemic wavelength. The reference wavelength Åis taken as the midpoint of the two calibration lines, Å $= (6532.90 + 7173.91)/2 =$ 6852.9Å which falls near the center of the spectral range. The velocity at the reference stripe is 0.00 km s^{-1}, confirming the formula is correctly zeroed. The rotation curve spans −277.0 to +247.7 km s^{-1} across the measured radius, shown in Fig. 3. A fifth-order polynomial fit to the rotation curve yields residuals with a standard deviation of 13.75 km s^{-1}, which is adopted as σ_v throughout the error propagation.

3.5. Galactic Radius and Enclosed Mass

The physical radius at each stripe was computed from the CCD plate scale of 0.468 arcsec pixel^{-1}, a galaxy distance of 150.7 Mpc derived from the systemic redshift of 10,100 km s^{-1} and $H_0 = 67$ km s^{-1} Mpc^{-1}, and the small-angle approximation. The measured radius spans ±20.52 kpc, consistent with the earlier result of Fraustro (2019) and within the expected range for a spiral galaxy of this type (Catinella et al. 2005). The stripe positional uncertainty of ±0.5 pixels corresponds to a radius uncertainty of ±0.171 kpc.

Enclosed mass was computed via the Kepler relation $M(r) = v^2 r/G$, with uncertainties propagated as described in Section 2.5. The enclosed mass ranges from 0 at the nucleus to $3.33\times10^{11}\,M_\odot$ at the outermost measured radius, with mass uncertainties ranging up to $3.48\times10^{10}\,M_\odot$. This is a significant improvement from the apparent negative mass that was shown in Fraustro (2019).[6] The enclosed mass profile with 1σ uncertainties is shown in Fig. 4.

3.6. Dark Matter Detection

To isolate the dark matter component, the dynamically inferred enclosed mass was compared against a luminous mass profile derived from the galaxy's own flux distribution, normalized in the inner region ($|r| \le 5.13$ kpc) where baryonic matter is expected to dominate and the rotation curve is rising. The mass-to-light ratio in this normalization region is 2.581×10^{37} kg count^{-1}.

In the outer 25% of the measured radial extent (beyond ≈ 15.4 kpc) the dynamical mass exceeds the luminous mass pro-

[6] A discovery which would have earned his second Nobel prize.

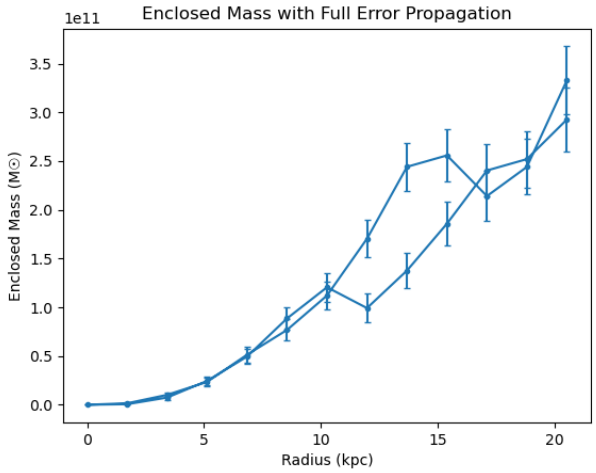

Fig. 4: Enclosed mass profile of UGC 9039 with error propagation.

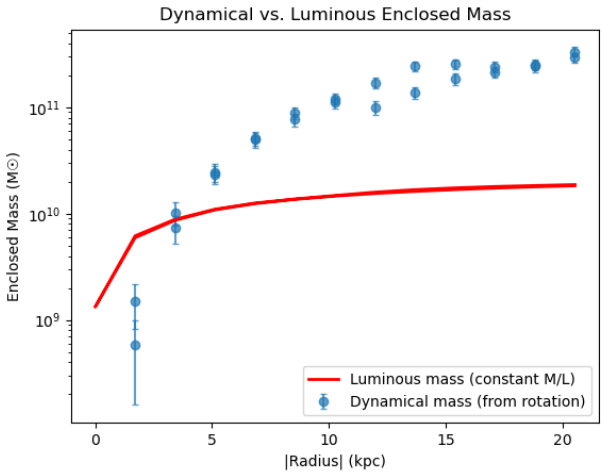

Fig. 5: Dynamical mass vs. luminous mass profile of UGC 9039.

file by $(2.37 \pm 0.12) \times 10^{11}\ M_\odot$, corresponding to a dark matter fraction of 92.8% at large radius. The detection significance is 20.0σ. The mass comparison is shown in Fig. 5.

The flat rotation curve at large radius seen in Fig. 3,the feature that Fraustro (2019) correctly identified as suggestive of non-Keplerian dynamics, but could not quantify, is now resolved at sub-pixel precision and confirmed at high significance.

The conclusion that Fraustro (2019) declined to state definitively, citing the absence of error analysis, is here stated: UGC 9039 contains dark matter.

4. Conclusions

UGC 9039 contains dark matter. The rotation curve is recovered cleanly across ±20 kpc, and the dark matter content of the galaxy is detected at 20σ significance.

No new physics is reported here. The galaxy was identified as a dark matter-bearing rotator long before the data analyzed here were taken in 2002 (Catinella et al. 2005). The contribution of the present work is the recovery of a detection that was, in 2019, sitting inside the exact same dataset, awaiting only an analysis pipeline that did what its author intended, and an author willing to write one.

The data did not change. The conceptual approach did not change. UGC 9039 certainly did not change. What changed was the execution. The cross-correlation routine was a few lines of Gaussian fitting away from working, the velocity formula was a single arithmetic operation away from being right, and the error analysis was an afternoon of work away from existing. The 2019 author knew, at some level, that all three of these were true. None of these fixes required new theoretical insight or techniques unavailable to an undergraduate student; they simply required time, attention, and a willingness to see what the code was actually doing, rather than what it was meant to do.

In the years since the original submission, the author has had occasion to reflect on the difference between code that works and code that is correct. The original pipeline ran to completion. It produced figures and numbers. None of these facts had any bearing on whether the numbers were right, and the smooth-looking velocity curve actively obscured the fact that they were not. The present reanalysis is offered, in part, as evidence for the proposition that one should not trust a pipeline merely because it terminates.

The 2019 version of this analysis was written under extreme time pressure, fueled by caffeine and necessity, with the conviction that the next compounding error would be the last. The 2026 version is written by a professional developer who has acquired a deep, consequential appreciation for the value of writing code slowly. Both versions of the author would benefit from being told, by someone with sufficient authority to be heeded, to take a break and come back tomorrow. Neither would have listened.

However, the original work succeeded in one respect that mattered most: the author declined to make a false claim to mask a failing pipeline. The original error analysis consisted of an explicit acknowledgment that there was no error analysis. The numbers were wrong, the physics was mangled, but the scientific reasoning was sound.

The dataset is now closed. The galaxy rotates, the dark matter is present at 20σ, and the project that has lived in the back of the present author's mind for seven years is, at last, finished.

Acknowledgements. The author thanks Dr. J. Harrington – Joe – for the original assignment, the patience extended during its initial submission, the pedagogical insistence on careful error propagation that is here belatedly honored, and most importantly, for the remarkably kind recommendation letter in 2021 that helped launch the present author's career at STScI. Finally, the author thanks his former self for committing the reduced products to version control before things went entirely sideways—an act of unwarranted foresight to which the present work owes its existence.

References

Fraustro, J. 2019, Radial Velocity And Dark Matter Analysis of Spiral Galaxy UGC 9039
Catinella, B., Haynes, M. P., & Giovanelli, R. 2005, AJ, 130, 1037