By the latest estimate, the observable universe contains 200 billion galaxies. Astronomers wonder: Why so few?

By James E. Geach

Forget dark matter: even the supposedly normal matter of the universe is mysterious enough. Why does only a small fraction of it reside in galaxies? Where did the rest go?

The current best guess is that the bulk of the normal matter is trapped in giant gaseous filaments. This so-called warm-hot intergalactic medium, or WHIM, is hard to detect directly.

Galaxy formation is evidently rather inefficient. As material falls into a galaxy, the galaxy tends to shoot much of it straight back out again—a process known as feedback.

The atoms in your body have probably been cycled through intergalactic space. Indeed, galaxies and their contents are not fixed structures but the bright tips of a wider sea of gas.
IN THE THROES OF FORMATION, a galaxy like our Milky Way pulls in dense, cold gas (red streams) and also ejects hot gas (blue streams) back into intergalactic space. The galaxy ends up with only a small fraction of the raw material. The author and his colleagues generated this image using a state-of-the-art cosmological simulation code.
Honey, We Lost Half the Universe

FOR DECADES OBSERVERS HAVE BEEN PIECING TOGETHER A TIMELINE OF COSMIC HISTORY THAT DESCRIBES THE CONTENT OF THE UNIVERSE AT DIFFERENT STAGES IN ITS EVOLUTION. IN THE PROCESS, IT HAS GRADUALLY BEEN REVEALED THAT THE UNIVERSE IS DIFFERENT AT EACH STAGE.

The initial amount of baryonic matter is actually fairly straightforward to estimate. This information is encoded into the Planck space observatory detect small fluctuations in the temperature of this radiation, and the distribution of these fluctuations reflect the baryon density of the universe at a time when galaxies were yet to form. An independent check comes from measurements of the abundances of helium, deuterium and lithium. These elements were synthesized in the first few minutes of the universe in relative amounts that depended on the overall quantity of baryonic matter. Both techniques imply that the total amount of baryonic matter should account for 4 percent of the mass of today's universe.

At the outset, all baryons took the form of a hot gas that filled space. In the regions where the initial matter density was high, gravity caused gas to clump into progressively denser clouds, which were the starting point for galaxy formation. Astronomers have detected this reservoir of gas in the early universe by analyzing the light detected from bright, distant quasars. What quasars are is not important for now; just consider them extremely bright lighthouses, acting as backlighting for the primordial gas floating around in intergalactic space. When a light ray from a quasar passes through a cloud of cold, neutral hydrogen, the gas absorbs some of the photons. Because the gas absorbs only photons of a certain energy, it imprints a telltale dip in the quasar’s spectrum at a very specific wavelength: what astronomers call an absorption line.

A ray of light from the quasar might pass through hundreds of such clouds on its journey through the universe, and each one can imprint an absorption line at a slightly different wavelength, depending on the cloud’s distance from the observer. By summing up the dips, we can calculate how many of the baryons were locked into these clouds. The result suggests that as late as five billion years after the big bang, or about nine billion years ago, the initial allotment of baryons could still be accounted for. Most were floating around in intergalactic space and had not yet collapsed into luminous galaxies [see “The Emptiest Places,” by Evan Scannapieco, Patrick Petitjean and Tom Broadhurst; SCIENTIFIC AMERICAN, October 2002].

In the intervening nine billion years, most of the galaxies we

How did all those galaxies come to be? This question inspired me to become an astronomer and has been the focus of my research career. Over the years my naive way of looking at galaxies has changed. To judge by their sheer numbers, nature appears to be quite good at producing galaxies. Not so. If you add up all the visible matter in galaxies today, you get only about a tenth of the total endowment created by the big bang. Where is the rest, and why did it not end up in galaxies? These are two of the biggest puzzles in astronomy today.

This missing matter is different from dark matter and dark energy. These are substances of unknown composition that together amount to 96 percent of the total mass of the cosmos. In this case of numbers of galaxies, the trouble is with the 4 percent that was supposed to be well understood. This slice of the universe is normal matter, made of the same stuff as our bodies and everything around us—primarily baryons, the class of particle that includes protons and neutrons. For the majority of it to go missing is a mystery inside a mystery. Not only is most of the matter in the universe dark and unexplained, but of the small sliver that is normal, only a fraction is accounted for.

Another way to put it is that the process of galaxy formation must be inefficient. It is as if a farmer sowed an entire field of seeds and only one of every 10 germinated. Astronomers have struggled for years to explain how that could be. The emerging answer requires us to revise our notions not only of how galaxies form but of what a “galaxy” even is. One expects not to fathom exotic types of matter; it is rather more disturbing to learn that we still do not grasp the mundane sort.

The initial amount of baryonic matter is actually fairly straightforward to estimate. This information is encoded into the relic radiation from the big bang: the cosmic microwave background radiation. State-of-the-art experiments such as the Wilkinson Microwave Anisotropy Probe and the Hubble Ultra Deep Field, captures some 10,000 galaxies in an area about 1/100th the size of the full moon. Scaled up to the whole sky, such a density implies a total of 200 billion or so galaxies. And those are just the most luminous ones; the true number is probably much larger.

Scientific American, May 2011
see today formed, wrought from that vast reservoir of primordial hydrogen. Once inside galaxies, baryons were reprocessed and took on various guises: stars, stellar remnants, neutral gas (both atomic and molecular), ionized gas, dust, planets, people. We can audit the mass of baryons in these different forms by measuring their emissions across the electromagnetic spectrum. For example, visible and near-infrared light reveal the mass of stars; a distinctive radio emission line signals the amount of neutral atomic hydrogen; infrared light betrays interstellar dust. In these ways, astronomers have taken a census of all the different phases of baryons in all the galaxies around us, and here is where the discrepancy arises: the total accounts for only 10 percent of the initial inventory of baryons that were present in the early universe. Presumably they did not simply vanish; they linger in the vast spaces between the galaxies. But why can we not see them?

ON A WHIM

Astronomers know where to find some of these intergalactic baryons. Dense swarms of galaxies called clusters are filled with a diffuse ionized gas, or plasma. The intense gravitational field of the cluster whips ions to high speed, giving the plasma a temperature of hundreds of millions of kelvins, enough to make it glow with x-rays. Space telescopes such as XMM-Newton and Chandra routinely detect clusters of galaxies by means of this x-ray emission. But clusters are rare, so the gas within them accounts for only another 4 percent of the baryons. When we add up all the baryons we can see in galaxies, clusters and elsewhere in intergalactic space, they account for roughly half the total, leaving the equivalent of at least 500 billion galaxies waiting to be found.

To balance the books, Renyue Cen and Jeremiah P. Ostriker, both at Princeton University, Romeel Davé of the University of Arizona, and their collaborators conjectured a decade ago that the missing baryons are out there but have evolved into a phase that is hard to detect. The properties of this elusive component are related to the way that cosmic matter of all types, both mundane and exotic, has developed what astronomers refer to as large-scale structure.

Under the influence of gravity, dark matter has pulled itself into a vast skeletal network that interlaces the universe. Clusters are actually just the high-density nodes of this cosmic web. Outside clusters the majority of galaxies congregate in lower-density groups or line up in long filaments. Intergalactic gas is gravitationally attracted to filaments, and as it falls in, simulations suggest, it gets heated by shock waves to temperatures ranging from 100,000 kelvins to tens of millions of kelvins. That sounds hot but is tepid by the standards of intracluster gas. It is toasty enough to remain highly ionized but too cool to blaze in x-rays.

Astronomers think they may have found where the bulk of the normal matter in the universe lurks: not in galaxies but in a form of intergalactic gas (mostly hydrogen) called the warm-hot intergalactic medium, or WHIM. The name connotes that the gas is less than blazingly hot and, consequently, glows too feebly to see directly. Looking in the interstices of a giant filament of galaxies called the Sculptor Wall, astronomers saw, in essence, the WHIM's shadow: the gas absorbed x-rays from a background object at a distinctive wavelength.
Cen, Ostriker and Davé dubbed this material the warm-hot intergalactic medium, or WHIM. If we could empirically confirm its presence and extent, we might be able to pin down the location and condition of the missing baryons.

The most promising way to detect the WHIM is to look for trace constituents such as ionized oxygen or nitrogen, which absorb ultraviolet light or x-rays of distinctive wavelengths. In fact, astronomers can apply the same absorption-line technique we used for the census of cold hydrogen clouds in the early universe; namely, we can look for dips in the spectra of quasars that back-light the WHIM. We have already had tantalizing glimpses. In the ultraviolet, the Hubble Space Telescope and the now defunct Far Ultraviolet Spectroscopic Explorer (FUSE) have detected absorption by strongly ionized oxygen. The first hints came just over a decade ago, when the concept of the WHIM was still novel. Todd M. Tripp, now at the University of Massachusetts Amherst, and Blair D. Savage of the University of Wisconsin-Madison detected ionized oxygen absorption in the far-ultraviolet spectrum of the quasar PG 0953+415. More observations have followed over the past decade, helped by improvements in detector technology and instrumentation, the most recent being the installation of the Cosmic Origins Spectrograph on Hubble. Although systems with strongly ionized oxygen appear to be plentiful, this ion traces only the relatively cool part of the WHIM. To trace the more abundant hotter gas, we must search for absorption by even more highly ionized species.

Taotao Fang of the University of California, Irvine, and his collaborators, have used the x-ray telescopes Chandra and XMM-Newton to peek into the interstices of the Sculptor Wall, a vast string of galaxies in the local universe—perfect WHIM-hunting territory. They have found absorption by oxygen that is so strongly ionized that it has lost almost all its electrons. The team estimated that the overall baryon density in this WHIM component agreed with cosmological simulations.

Though encouraging, these observations only scratch the surface. The observations are hard: the WHIM signal is weak, and we are generally working at the technical limits of the instrumentation. Even when we do detect absorption, we have to make many assumptions about the makeup of the gas to extrapolate the wider WHIM properties. More important, the absorption-line technique relies on fortuitously placed quasars. Quasars are rare, bright ones even more so, which makes WHIM hunting somewhat of a lottery. Nevertheless, we think we know where the missing baryons are and how they might be detected. Many astronomers are now engaged in efforts to map out the WHIM properly.

**THE BATTLE FOR BARYONS**

The existence of the WHIM goes some way to explaining why galaxy formation is so inefficient. The evolution of large-scale structure made intergalactic gas too tenuous and hot to accumulate into the cool, dense pools required for galaxy formation. Obviously, though, some of the baryons did manage to turn into galaxies, or else we would not be here.

Another thing is clear, too: galaxy formation used to be much more efficient. About eight billion years ago the average birth rate of stars was 10 to 20 times higher than it is today. Most of the galaxies we see today took shape then. To account for why galaxy formation slackened as sharply as it has, astronomers have had to rethink our basic models of how galaxies are born.

In principle, the recipe for a galaxy is quite simple. In a model pioneered in the 1990s by Simon D. M. White of the Max Planck Institute for Astrophysics in Garching, Germany, and Carlos S. Frenk of Durham University in England, galaxies grow within massive clumps of dark matter, termed halos, whose gravitational attraction sucks in surrounding gas like water going down a plug hole. In this model, some of the gas gets heated by shock waves as it plows into the halo and then cools by emitting radiation, allowing it to agglomerate into a cohesive body. Once within the galaxy, the gas can cool further and collapse into clouds of molecular hydrogen. Under gravitational contraction, these clouds can eventually reach the density required to make stars. Bigger galaxies can grow through the mergers of smaller ones.

White and Frenk recognized that their model could not be the entire story, however. For instance, not all of the gas flowing into galaxies would be shock-heated to high temperatures. But the basic picture of gas accretion within dark halos gave astronomers a solid framework within which to understand the principles of galaxy formation. The field has blossomed in the past 20 years. Theorists have explored the physics of gas flow in ever more detail, refining the original model. Recent high-resolution computer models of the thermodynamic evolution of gas in cosmological simulations suggests that some of the gas flowing into young galactic disks in the early universe does so in streams that are relatively cold (10,000 to 100,000 kelvins) and narrow (a few thousand light-years across). These cold flows appear to penetrate the hotter halo gas and directly feed the galaxies.

No one has yet seen this process in action: the detailed physics of the accretion of gas onto galaxies is complicated, and different simulations predict slightly different things. These caveats aside, astronomers now accept that all galaxies build up from the gravitational accumulation of primordial gas, be it gas that heats up and cools down or gas that never heats up at all.

The trouble with this model is that the flow of gas into galaxies cannot go on unabated. If it did, galaxies would grow into monsters, and we know they do not: galaxies today come in only a limited range of masses. Early models seemed to reproduce the observed range of galactic masses pretty well, but in retrospect they worked only because astronomers were using a value for the overall baryon density that was about half the
present value. As new measurements of the baryon fraction revised the value upward, theorists fed this information into simulations and realized that their model universes were plagued by a serious overabundance of massive galaxies that are not seen in nature.

Another problem is that models predict a profusion of smallish dark matter clumps that agglomerate into progressively larger bodies. Real galaxies do not follow this pattern. Observers do not see nearly as many small galaxies as the models predict, and the most massive galaxies appear to have formed quickly and efficiently, rather than through the gradual assembly of smaller pieces.

The models were clearly missing a critical ingredient. Something must be regulating the cooling of gas and the formation of stars in galaxies. The process has made small galaxies inefficient at forming stars and limited the size of massive galaxies. Theorists began considering a variety of additional physical processes that would provide this regulation. Known collectively as galactic feedback, these processes can counter, or reverse, the gravitational collapse of gas into galaxies and thus limit the number of stars that can form. They include supernovae explosions, stars' ultraviolet radiation and outflows, and the tremendous energy released during the growth of the supermassive black holes that lurk in the core of all massive galaxies [see "Black Hole Blowback," by Wallace Tucker, Harvey Tananbaum and Andrew Fabian; SCIENTIFIC AMERICAN, March 2007]. In the most massive galaxies, black holes are probably the most dominant feedback mechanism; in lower-mass systems, supernovae and stellar winds are more important.

What all these processes have in common is that they inject energy back into the surrounding medium. In this way, galaxies can choke off the inward flow of material, prevent gas that has already accumulated from forming stars or, in extreme cases, eject baryons back into intergalactic space. Simulations that take feedback into account do a much better job of reproducing the observed variety of galaxies. Not only does feedback play a critical role in tuning the evolution of galaxies, it can also re供应, reheat and enrich the WHIM. Through a continuous process of cooling and heating, baryons are cycled between intergalactic space and the stars and gas within galaxies. Galaxy growth is determined by a delicate balance of power that has tipped one way or another over cosmic history. Understanding this battle for baryons revamps our view of galaxy formation.

**BLOBOLOGY**

The study of gas cooling and feedback has been a major focus of astrophysics over the past decade. Without empirical data, we have no way of testing the models. Cold material flowing into galaxies in the early universe should give itself away by the diffuse glow that hydrogen emits as it cools. Feedback can be inferred from the bright infrared emission from intense star formation and x-ray or radio emission from the environs of a supermassive black hole. We need to catch both these processes in the act. Recently we may have done just that.

About a decade ago Charles Steidel of the California Institute of Technology and his collaborators discovered a new class of object that appears to tick the boxes for the observational signature of cooling: the Lyman-alpha blob. Never let it be said that astronomers are uptight with their nomenclature; “blob” really is the technical term. Lyman-alpha refers to one of the specific
Most people think of galaxies as stately structures: giant balls or majestic pinwheels of stars floating in the emptiness of deep space. Astronomers are finding, however, that galaxies are fluid systems that actively exchange material with their surroundings. Normal matter cycles through galaxies, and at any given moment most of it resides in intergalactic space. This simulation shows a galaxy akin to our Milky Way as it would have appeared 10 billion years ago and today.

**Young, Chaotic Galaxy**

Early on, complex and chaotic currents of gas move in both directions: gravity pulls in dense gas, fueling the unstable galactic core (red/white streamlines), and feedback processes eject hot gas (blue/white streamlines).

**Established Galaxy**

The inflow and outflow diminish, and much of the cold gas is trapped in a rotating disk. Surrounding the disk is a hot, gaseous halo, which continues to swap material with intergalactic space. Gas can rain back down onto the disk, replenishing it.
frequencies of light that hydrogen gas emits. These blobs appear to be glowing clouds up to 300,000 light-years across—for bigger than our Milky Way galaxy—making them among the largest luminous objects in the early universe. Astronomers have since discovered scores of them. The observed Lyman-alpha glow bears an uncanny resemblance to theoretical predictions of the radiative signature of cold gas flowing into young galaxies.

On the other hand, many other astrophysical processes could cause Lyman-alpha emission, too. For instance, ultraviolet light or a galactic-scale wind could pump energy into the blobs and cause them to glow. Using Chandra, my colleagues and I have shown that many blobs contain galaxies with actively growing black holes that shine brightly in x-rays. Often this activity is accompanied by intense star formation, revealed by its infrared emission from the obscuring layers of dust that blanket stellar birthing grounds. We have calculated that the energy released by these processes is more than enough to power the Lyman-alpha emission. So perhaps the blobs’ glow is caused not by cooling, as many think, but by heating.

Rather than making things clearer, these blobs have muddied the waters somewhat. But that is what excites me about this field—it would not be science if we knew all the answers. We must now devise and conduct new observations to try to unveil what is really going on. Either way, though, blobs are precisely the type of object that could fill in some of the major gaps in our understanding about the origin of galaxies.

The observation that the intergalactic medium surrounding young, active galaxies gets swamped by radiation could help resolve another problem with galaxy-formation models. Very high resolution simulations of dark matter predict that galaxies such as the Milky Way should be accompanied by thousands of lower-mass dwarf galaxies buzzing around them like bees around a hive. Although the Milky Way does have a few dwarf companions, they are far fewer than the simulations predict.

One solution could be that the dwarf galaxies did form in the early universe, but their parent galaxy blasted them with radiation and winds. The barrage stripped away any baryons the dwarfs had managed to accumulate, leaving only barren clumps of dark matter that have skulked on the outskirts of the parent galaxy ever since. Larger galaxies reach a truce in the battle for baryons, but smaller ones lose the battle entirely.

**WHAT IS A GALAXY?**

**Perhaps the most exhilarating experience a scientist can have is the feeling of a sea change in one’s perspective on the world. For me, that came when I had to reevaluate what I thought of as “a galaxy.” Traditionally we think of luminous galaxies as isolated and discrete island universes, as German philosopher Immanuel Kant put it. In some sense, that is clearly true. But the bright galactic islands of light are just the visible tips of a much wider, but still elusive, sea of baryonic matter. This material pervades the universe, distributed within and shaped by a vast, underlying dark architecture, continuously evolving through gravity.**

All those baryons started off in the same state: a hot, pristine gas that rapidly formed the basic elements of hydrogen and helium, along with small amounts of deuterium and lithium. What we think of as galaxies formed from this raw material, pulled into dense concentration by gravity. But these structures are not fixed groups of baryons. Material moves among them as part of a vast cycle that has been in operation since the big bang. The competing influences of gravity and feedback cause gas to cool onto, and later get ejected from, galaxies. Recent computer simulations by Rob Crain of the Swinburne University of Technology in Melbourne, Benjamin Oppenheimer of Leiden University in the Netherlands and their collaborators suggest that up to half of the baryons currently locked into galaxies in the local universe have cycled through the intergalactic medium at least once and often many times. The baryons that make up your body have participated in this cycle for nearly 14 billion years; the matter within your fingernail could have formed in stars in other galaxies and then spent billions of years exiled in intergalactic space before coming to rest in our solar system. You are just an ephemeral phase, a brief host, to this rare substance we call “normal.”

This concept of baryon cycling underpins the emerging view of galaxy evolution. The big picture you should have in mind is that galaxy evolution is just a small component of the large-scale evolution of the intergalactic medium. The baryonic universe is predominantly gaseous, not galactic. The intergalactic medium is a battleground of forces, and amid this maelstrom, galaxies form.

Galaxies are just one processing stage in a cycle that is continuously shifting baryons from one phase to the next, and at any one time, most of the baryons in the universe are not inside galaxies.

Sentimentally, perhaps, we hold galaxies in a special regard: the Milky Way is our cosmic habitat, a brilliant, vast, complex home within the darkness. From an anthropic viewpoint, we just happen to be lucky enough to exist at a time when the baryons that make up Earth and everything on it have taken on a cold, stable form.

That will not always be the case. The death of the sun around five billion years from now will incinerate the inner planets, evaporate the outer ones and gradually disperse the resulting detritus of heavy elements back into the interstellar medium. Unless humans manage to cheat the cycle by developing the technical capability to escape the confines of the solar system, the ashes of every material thing on Earth are fated to be returned, enriched, to the cosmos. And so the cycle continues.